Nous paradigmes de la contribució paterna al desenvolupament embrionari primerenc mitjançant l'anàlisi de l'espermatozoide humà amb tècniques d'alt rendiment

Ada Soler-Ventura

UNIVERSITAT DE BARCELONA

Nous paradigmes de la contribució paterna al desenvolupament embrionari primerenc mitjançant l'anàlisi de l'espermatozoide humà amb tècniques d'alt rendiment

Ada Soler-Ventura

Tesi Doctoral

Barcelona, 2021

Tesi doctoral

Nous paradigmes de la contribució paterna al desenvolupament embrionari primerenc mitjançant l'anàlisi de l'espermatozoide humà amb tècniques d'alt rendiment

Memòria de Tesi Doctoral presentada per

Ada Soler-Ventura

per optar al grau de

Doctora per la Universitat de Barcelona

en el

Programa de Doctorat en Medicina i Recerca Translacional

Treball desenvolupat sota la direcció de

Dr. Rafael Oliva Virgili

Dra. Meritxell Jodar Bifet

Grup d'investigació en Biologia Molecular de la Reproducció i el Desenvolupament, Unitat de Genètica, Departament de Biomedicina, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona, i Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS)

Director	Directora	Autora
Dr. Rafael Oliva Virgili	Dr. Meritxell Jodar Bifet	Ada Soler-Ventura

Als petits, als grans, als que hi són, als que no hi són,
A tots els que estimo,
A tu.

"Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less"
— **Marie Curie**

"Eppur si muove"
— **Galileo Galilei**

Aquesta Tesi Doctoral és el resultat de tota la feina feta durant els darrers anys, però no hagués estat possible sense l'ajuda, dedicació, recolzament i estima de totes les persones que han estat al meu costat i que m'han fet créixer com a persona i com a científica.

Primer de tot vull agrair de manera molt especial als meus directors de Tesi, el Dr. Rafael Oliva i a la Dra. Meritxell Jodar. **Rafael**, moltes gràcies per deixar-me formar part de la petita gran família que és el Grup d'Investigació en Biologia Molecular de la Reproducció i el Desenvolupament. Gràcies per l'ajuda, l'aprenentatge i guia que m'has brindat els últims anys. Ets un referent per mi. M'agradaria recordar una frase que em vas dir en un *Lab Meeting* en què els meus resultats no eren els esperats i que reflecteix un dels grans missatges que m'has transmès: "Això és ciència i la ciència és veritat". M'emporto l'aprenentatge d'un gran científic i mentor com tu. **Meri**, què dir-te que no sàpigues ja. Gràcies per ser-hi i estar-hi, per haver-me ensenyat tot el que sé amb paciència infinita. Gràcies per fer de directora de tesi, d'amiga, de profe, donar-me energia quan creia que ja no en tenia, donar-me canya quan ho mereixia o una ràfega de confiança en mi mateixa quan la necessitava. Gràcies per confiar en mi i donar-me l'oportunitat d'entrar al famós i fantàstic món de Monstruos RNA, que sempre perilla per les RNases. Gràcies per compartir amb mi els projectes i hipòtesis més crazy, des de buscar mutacions de proteïna a DNA, passant per teories boges fins a extreure RNA d'espermatozoide d'escarabat! Gràcies per despertar-me la curiositat, el rigor i la capacitat d'anàlisi. Gràcies a tu avui sóc la científica que sóc i sé, que tot el que aconsegueixi en un futur, en part, serà gràcies a tu. Les paraules se'm queden curtes si et dic que ets i seràs un referent i una mentora per mi.

Vull fer un agraïment especial als meus companys de laboratori. A **l'Orleigh**, en **Claudio** i la **Montse** amb els que vaig coincidir poc, però que em van ajudar molt al principi. Als membres actuals, que més que companys s'han convertit en amics (Us estimo!). **Ferran** (Dr. Barranquilla), vam començar junts aquesta gran aventura. Gràcies per ensenyar-me, recolzar-me, escoltar-me i ajudar-me sempre que ho he necessitat i estar al meu costat durant tots aquests anys a les dures i a les madures, a les 10 del matí o a les 3 de la matinada. T'admiro molt i et desitjo el millor en aquesta nova etapa que has començat a l'altra punta del món. **Alberto** (de la Iglesi), que decir que no sepas? Gracias por hacerme esta portada tan chula y no desesperarte por mis cambios tontos. Compañero de dramas, de cafés, de alegrías y recortes (Que Monopoly más chulo hahaha). Por siempre ser Team Ladri conmigo. Por escucharme, animarme, darme energía, curiosidad y momentos de relax. Por siempre estar tanto en lo personal como en lo profesional. Por los cafés

mañaneros y las cervezas que siempre apetecen y las que vendrán. Eres una gran persona, gran amigo y mejor científico, te vas a comer el mundo. **Judit**, gràcies per ensenyar-me tantes i tantes coses, per compartir projectes (gordos) i per fer-me millor científica i persona. Hem compartit moments inoblidables (vam anar a viure en parella igual i ens vam casar amb tres mesos de diferència!). Gràcies per ajudar-me a fer càlculs quan em perdia, gràcies per ajudar-me a dissenyar experiments, a ser rigorosa, detallista i crítica. Vals moltíssim i aconseguiràs totes les metes que et proposis! **Jose**, gracias por siempre estar cuando lo necesito, por tener siempre esa sonrisa y disponibilidad. Mucha suerte en esta etapa final, será duro, pero seguro que podrás con todo! També vull agrair a tots els estudiants que han passat pel laboratori durant la meva tesi. **David**, moltes gràcies pels cafès, les converses, l'ajuda amb la bioinformàtica, pujar dades al browser, scripts de R i els programes que em vas fer perquè fos més fàcil tot. Gràcies! **Ainitze**, muchas gracias por tu ayuda, por formar parte de Monstruos RNA. Vales un montón y llegarás muy lejos. Eskerrik asko! També a l'**Arnau**, la **Paula**, la **Marta,** la **Tosca** i en **Lukas**. Moltes gràcies a tots pel suport i l'ajuda! Estic segura que tindreu un futur brillant!

Tota la recerca que he realitzat al llarg de la Tesi no hagués pogut ser sense la participació de tots els membres clínics del grup, que són imprescindibles i ens donen una perspectiva diferent. Al Dr. **Josep Lluís Ballescà**, Dra. **Salvadora Cívico**, Dra. **Dolors Manau**, Dr. **Juan Manuel Corral**, Dra. **Marta Guimerà,** Dr. **Josep Maria Calafell** i a les tècnics que sempre ens donen un cop de mà, la **Raquel** i **l'Alícia**.

També vull agrair a les persones amb les que he col·laborat al llarg de la Tesi i de tots aquests anys. Als membres de la Clínica Eugin, en especial el Dr. **Marc Torra-Massana**, per haver compartit molts treballs de màster i un projecte de tesi. Espero que tot et vagi bé! El Dr. **Giueseppe Grande**, per haver portat una part de Roma al laboratori!

M'agradaria agrair a tots els integrants del departament de Biomedicina, que sempre que poden ens donen un cop de mà, Dr. **Arcadi Gual**, Dr. **Xavier Gasull**, Dr. **David Soto**, Dra. **Núria Comas**, Dra. **Ester Gratacós**, Dra. **Jovita Mezquita**, Dr. **Belén Mezquita**, **Montse Pau**, **Alba**, **Fede** i segur que me'n deixo. Moltes gràcies per la vostra ajuda i per la infinitat de vegades que m'heu donat cèl·lules per fer proves! També vull agrair a la **Núria Román**, secretària del departament, per sempre ser-hi, aclarir-me dubtes burocràtics i ajudar-me en tot el que necessités.

Una gran part de la meva Tesi ha estat possible, també, gràcies a l'ajuda incondicional del personal de la unitat de proteòmica dels Centrs Científics i Tecnològics de la UB (CCiTUB). Al Dr. **Josep Maria Estanyol** i la Dra. **Maria José Fidalgo**, per tot i pujar mil vegades a aclarir dubtes i discutir resultats, sempre m'ho han resolt i m'han ajudat en tot el possible. També, al Mass Spectrometry & Proteomics Unit de l'IRB, per endinsar-nos al món del *top-down* MS i ajudar-nos a posar a punt, entendre i interpretar els resultats de les protamines i veure que l'espermatozoide és totalment diferent a una cèl·lula convencional. He après molt de tots vosaltres, moltes gràcies per tot: Dra. **Marta Vilaseca**, Dra. **Marina Gay**, Dra. **Laura Villareal**, Dra. **Mar Vilanova** i Dr. **Gian Arauz-Garofalo**.

També voldria agrair a tot l'equip del Biobanc HCB-IDIBAPS: **Aina, Teresa, Èlia, Tere, Roser, Patri, Míriam, Regina, Mònica, Laura, Amanda, Sara, Vero** i tots els estudiants de pràctiques (**Víctor, Tinta, Mireia, Emma, Marina i Andrés**). Aquest any treballant amb tots vosaltres ha estat fantàstic! Gràcies per sempre interessar-vos i preguntar-me per la Tesi i donar-me ànims (molt necessaris a vegades). En especial vull agrair a la Dr. **Teresa Botta.** Gràcies per endinsar-me al món dels biobancs i gràcies per recolzar-me, donar-me coratge i posar-me totes les facilitats per poder combinar la feina amb la Tesi!

Vull agrair als meus amics de Barcelona i de Besalú. Gràcies per distreure'm, preguntar-me, recolzar-me i escoltar-me. En especial vull agrair a en **Marc**, per estar sempre al meu costat des del principi del màster en els bons i no tant bons moments i en breus, com vam dir, ens esperen 25 cerveses per compartir! Espero que ara que estàs començant la tesi comptis amb mi pel que necessitis i n'estic segura que ho faràs genial! A la **Vila** (i a la meva fillola **Elna**), gràcies per estar sempre al meu costat, per donar-me forces quan no en tinc i escoltar-me encara que no acabis d'entendre el que t'explico. Gràcies per no dubtar mai de mi, sentir-te orgullosa (i dir-m'ho sovint) i recolzar-me durant tota aquesta gran aventura. TSR!

Gràcies a tots els membres de la meva fantàstica família. Gràcies **Baba** per sempre estar orgullosa de mi. Gràcies **Carme, Pili** i **Josep** per sempre ser-hi i recolzar-me. Gràcies **Luisa, Juan, Michel** i **Gael** per sempre preocupar-vos per mi. Gràcies **Xicu**, **Muntsa** i **Laura** per compartir tants moments i ajudar-me al llarg d'aquests anys a Hercegovina, 14, sempre us estaré agraïda. Gràcies **Andreu** i Lea per ajudar-me a desconnectar, per interessar-vos sempre per la Tesi i per mi. Gràcies a les petites de casa (**Anna** i **Ona**), que saben que la tieta "és científica" i viu a Barcelona, em feu ser millor persona cada dia. Gràcies **Papa, Mama** i **Lea** per no dubtar mai de mi i donar-me

forces quan no en tenia. Per recolzar-me, encoratjar-me, empoderar-me i creure en mi quan sovint jo no hi creia. Gràcies també Papa, per la paciència infinita revisant-te la Tesi i al llarg d'aquests anys fent-me preguntes de tots els experiments (encara que no els entenguessis). Gràcies **Vane** per haver-me recolzat sempre, sé que siguis on siguis deus estar molt orgullosa de mi. Gràcies a tots i cadascun dels membres de la meva família, que sempre hi sou i mai falleu. No podria desitjar millor família que tots vosaltres! Us estimo!

I finalment tu, **Juan**, company d'aventures, company de vida i des de fa relativament poc, marit. Què dir-te que no t'hagi dit al llarg de tots aquests anys? Aquesta Tesi en part també és teva. Gràcies per recolzar-me, consolar-me, eixugar-me les llàgrimes quan creia que no podia més, per confiar en mi i el meu potencial, per fer-me la vida més fàcil (i muntar-me un "xiringuito" informàtic a casa), per donar-me xocolata en els moments difícils, cervesa en els mals i bons dies, menjar bo per animar-me i amor en tot moment. Gràcies per creure en mi quan ni jo hi creia, gràcies per ser-hi i estar-hi sempre. Gràcies per fer-me creure que seré tot el que em proposi, que valc molt i que no haig de tenir por a res. T'estimo Xux!

Gràcies de tot cor a tots i cadascun de vosaltres que heu format part d'aquesta aventura!

Suport econòmic:
Aquesta Tesi Doctoral s'ha dut a terme gràcies al suport econòmic dels projectes d'investigació PI16/00346, PI13/00699 i fons FEDER "una manera de hacer Europa" del Ministerio de Economía y Competitividad; al projecte SERONO 13-015 de la Fundación Salud 2000; al projecte EUREP 2014 de EUGIN-UB concedits al Dr. Rafael Oliva.

0.

ÍNDEX

0. ÍNDEX

1.

INTRODUCCIÓ

1. INTRODUCCIÓ

1.1. Espermatogènesi

L'espermatogènesi és un procés complex de diferenciació cel·lular on ocorren grans canvis molt marcats a nivell cel·lular, genètic, funcional i estructural que resulten en l'espermatozoide madur (de Kretser *et al.*, 1998; Lin and Troyer, 2014). S'estima que tot el procés té una durada d'uns 74 dies en l'espècie humana, però pot variar de 42 a 76 dies i la temperatura òptima és a uns 32 ºC (de Kretser *et al.*, 1998; Misell *et al.*, 2006).

L'espermatogènesi té lloc al testicle, específicament als túbuls seminífers que són la unitat funcional i formen una xarxa de conductes complexa anomenada *rete testis* que pot arribar a mesurar uns 250 m per testicle. Es pot dividir en 3 etapes clarament diferenciades: mitosi, meiosi i espermiogènesi que culmina amb l'alliberació de l'espermatozoide al lumen del túbul seminífer. En els túbuls seminífers s'hi troben principalment dos tipus cel·lulars: les cèl·lules espermatogèniques i les cèl·lules no espermatogèniques. Dins les cèl·lules espermatogèniques trobem, en ordre de diferenciació i de la part més basal al lumen, les espermatogònies, els espermatòcits, les espermàtides i finalment els espermatozoides que són alliberats al lumen del túbul seminífer (Figura 1.1). Principalment, les cèl·lules no espermatogèniques són conformades per les cèl·lules de Sertoli que formen la barrera hemato-testicular gràcies a les unions fortes que formen i nodreixen i controlen la migració cap al lumen de les cèl·lules espermatogèniques (de Kretser *et al.*, 1998; Holdcraft and Braun, 2004). Entre els espais que separen els túbuls seminífers s'hi troba: (i) el teixit conjuntiu, format per cèl·lules peritubulars que s'encarreguen, mitjançant filaments contràctils d'actina i miosina, de propulsar el fluid testicular que conté els espermatozoides immòbils; (ii) capil·lars sanguinis; i, (iii) les cèl·lules de Leydig que sintetitzen andrògens, com la testosterona, i estan involucrades en la regulació hormonal de l'espermatogènesi (Veure apartat 1.1.2) (Holdcraft and Braun, 2004).

1.1.1. Fases de l'espermatogènesi

Durant el desenvolupament embrionari, es creen les cèl·lules primordials germinals (PGCs) a partir del teixit extra-embrionari. Entre la setmana 3-5 de desenvolupament migren cap al que serà el testicle i es diferencien a gonòcits, que quedaran aturats en fase Go i estaran mitòticament inactivats fins després del naixement. Just després del naixement i fins als 6 mesos de vida, els gonòcits es diferencien a espermatogònies, que restaran en estat quiescent fins a l'entrada de la pubertat, que és quan les espermatogònies comencen a diferenciar-se fins a

l'obtenció de l'espermatozoide i es mantindrà constat durant tota la vida fèrtil de l'individu (de Kretser *et al.*, 1998; Goldberg and Zirkin, 2018).

1.1.1.1. Mitosi

Les espermatogònies tenen dos papers principals que són l'autorenovació a través de la proliferació mitòtica, on hi ha fenòmens de replicació del DNA, transcripció i traducció que asseguren la disponibilitat d'espermatogònies al llarg de la vida reproductiva; i l'entrada en meiosis que permet la producció de l'espermatozoide haploide. S'han descrit tres tipus d'espermatogònies, les A de nucli fosc, les A de nucli pàl·lid i les B i, totes elles se solen situar a la part basal del túbul seminífer (Figura 1.1).

De manera consensuada, s'accepta el model de que les espermatogònies A de nucli fosc, que presenten una elevada condensació de la cromatina i una taxa baixa de replicació, romanen de forma quiescent i només es dupliquen per donar lloc a les espermatogònies A de nucli pàl·lid. Les espermatogònies A de nucli pàl·lid presenten una cromatina poc condensada i una elevada taxa de replicació, que permet la seva autorenovació i la formació de les espermatogònies B. Finalment, les espermatogònies B realitzen una divisió mitòtica i seguidament inicien la meiosis. Tot i així, les divisions són incompletes i les cèl·lules generades rauran interconnectades per ponts citoplasmàtics que es mantindran fins al final de l'etapa meiòtica (Figura 1.1) (de Kretser *et al.*, 1998).

1.1.1.2. Meiosi

La fase meiòtica s'inicia quan les espermatogònies B experimenten un augment del seu citoplasma, dupliquen el seu DNA i esdevenen espermatòcits primaris. En aquest moment, els espermatòcits primaris entren a la primera divisió meiòtica, que dura aproximadament 24 dies i és quan es dona la replicació del DNA, la condensació dels cromosomes i la recombinació genètica (Figura 1.1). Finalment, es redueix el contingut cromosòmic, degut a la migració dels cromosomes homòlegs a les cèl·lules filles. Al final d'aquesta etapa, les cèl·lules resultants són els espermatòcits secundaris, que són les primeres cèl·lules haploides. La segona divisió meiòtica sols dura unes 6 hores, no hi ha replicació del DNA i se separen les cromàtides germanes, esdevenint espermàtides rodones (Figura 1.1). A mesura que la meiosis avança, les cèl·lules en els diferents estadis van avançant al llarg del túbul des de la part més basal fins al lumen, en humans de manera helicoïdal. Finalment, de cada espermatogònia B en resulten 4 espermàtides rodones (Figura 1.1) (de Kretser *et al.*, 1998).

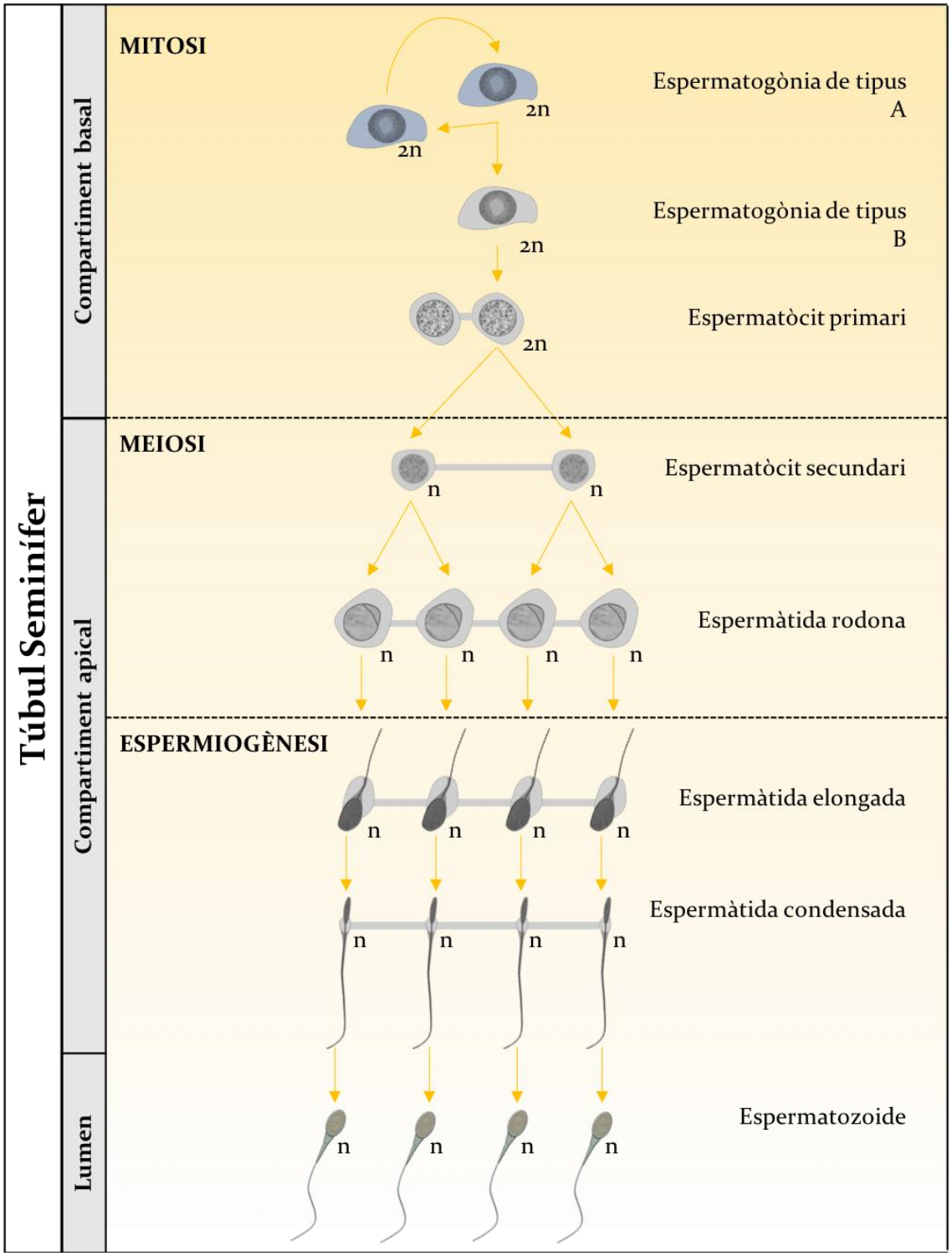


Figura 1.1. Procés de l'espermatogènesi humana. Les espermatogònies A de nucli fosc, es dupliquen per donar lloc a les espermatogònies A de nucli pàl·lid. Les espermatogònies A s'autorenoven i formen les espermatogònies B. Seguidament, les espermatogònies B realitzen una divisió mitòtica, inicien la meiosis i esdevenen espermatòcits primaris. Els espermatòcits primaris entren a la primera divisió meiòtica, donant lloc als espermatòcits secundaris. A continuació té lloc la segona divisió meiòtica, on els espermatòcits secundaris donen lloc a les espermàtides rodones. Finalment té lloc l'espermiogènesi, on les espermàtides rodones pateixen un dràstic procés de diferenciació cel·lular i esdevenen, primerament, espermàtides elongades i, finalment, espermatozoides. Basat en (Oliva and Castillo, 2011; Sharma and Agarwal, 2011).

1.1.1.3. Espermiogènesi

Durant l'etapa final de l'espermatogènesi, anomenada espermiogènesi, les espermàtides rodones pateixen un dràstic procés de diferenciació cel·lular on experimenten canvis citoplasmàtics, cel·lulars i morfològics, on primerament esdevenen espermàtides elongades i finalment espermatozoides (Figura 1.2). En els inicis d'aquesta etapa es produeix una gran activitat transcripcional i traduccional (Yan *et al.*, 2010). Bàsicament, es pot dividir en:

- **Condensació del nucli:** El nucli de l'espermàtida rodona migra cap a l'extrem de la cèl·lula i d'aquesta manera, la cèl·lula adquireix una estructura més allargada i esdevé espermàtida elongada. Així mateix, les histones que empaqueten la cromatina, són reemplaçades primerament per proteïnes de transició i finalment per protamines (Figura 1.2; veure apartat 1.4.3; veure annex 1). Les protamines empaqueten fortament el 85-95 % del DNA degut a la seva elevada càrrega positiva i provoquen la reducció de la mida del nucli i que la cèl·lula esdevingui transcripcionalment inerta (Cooper, 2005; Oliva, 2006; Oliva and Castillo, 2011; Jodar and Oliva, 2014).

- **Formació de l'acrosoma:** L'aparell de Golgi forma una gran vesícula a la part anterior de l'espermàtida, similar als lisosomes, que conté enzims hidrolítics essencials per l'entrada de l'espermatozoide a l'oòcit, s'aplana i acaba ocupant unes 2/3 parts del cap de l'espermatozoide (Figura 1.2). Durant la reacció acrosòmica (Veure apartat 1.2.2), la membrana interna i externa de l'acrosoma es fusionen i s'alliberen els enzims emmagatzemats que són necessaris per travessar la zona pel·lúcida i entrar a l'oòcit (Lin and Troyer, 2014).

- **Formació del coll, la peça intermèdia i el flagel:** Els centríols de l'espermàtida migren a la part posterior, la contrària a on es forma l'acrosoma i un dels centríols forma l'axonema del flagel. L'axonema està composat per nou parells de microtúbuls que envolten dos microtúbuls centrals (Estructura 9+2). Tanmateix els mitocondris també migren, concretament a la part proximal de l'axonema del flagel on es col·loquen helicoïdalment formant una beina i formen la peça intermèdia, la qual proveirà d'energia l'espermatozoide madur i permetrà el moviment flagel·lar (Figura 1.2) (Lin and Troyer, 2014).

- **Eliminació de la major part del citoplasma:** A mesura que la mida de la cèl·lula es va reduint, la major part del citoplasma incloent la maquinària traduccional és eliminat en forma de gota citoplasmàtica i fagocitat per les cèl·lules de Sertoli, fent que l'espermatozoide sigui traduccionalment silenciat (Figura 1.2) (Cooper, 2005; Oliva, 2006; Lin and Troyer, 2014).

Quan l'espermiogènesi finalitza, l'espermatozoide és alliberat al lumen del túbul seminífer (Figura 1.1), però li manca l'habilitat de moure's i fecundar. No serà fins a la seva maduració (Veure apartat 1.2.1) que té lloc en el seu trànsit al llarg de l'epidídim on l'espermatozoide adquirirà la seva motilitat i el potencial fertilitzant.

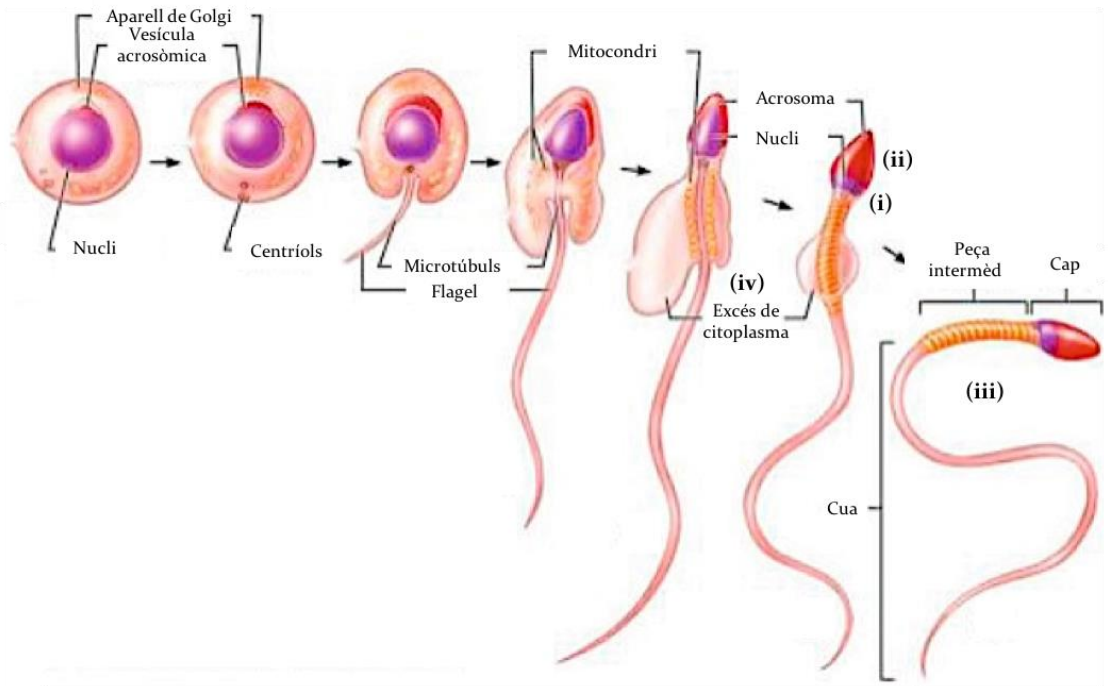


Figura 1. 2. Espermiogènesi. Els principals processos que tenen lloc durant l'espermiogènesi són: (i) La condensació del nucli; (ii) la formació de l'acrosoma; (iii) la formació del coll, la peça intermèdia i el flagel; i (iv) eliminació de la major part del citoplasma. Adaptat de (Lin and Troyer, 2014).

1.1.1.4. Estructures específiques de l'espermatozoide

L'espermatozoide té uns 60 µm de llargada. Es divideix principalment en el cap que té un diàmetre d'aproximadament 4-5 µm i la cua o flagel que presenta un diàmetre de 1-2 µm (Figura 1.3) (Holstein *et al.*, 2003).

- **Cap de l'espermatozoide:** Conté el nucli amb el DNA patern fortament empaquetat per les protamines, juntament amb proteïnes i RNAs (Veure apartat 1.4). Envoltant les dues terceres parts anteriors del nucli, es troba l'acrosoma que conté els enzims hidrolítics que permetran la penetració de l'espermatozoide a l'oòcit (Figura 1.3; veure apartat 1.2.2).

- **Cua de l'espermatozoide:** Conté el flagel que és l'encarregat de permetre la motilitat de l'espermatozoide. El flagel es divideix en quatre parts clarament diferenciades: el coll, la peça intermèdia, la peça principal i la peça terminal. Al coll de l'espermatozoide s'hi troben els centríols proximal i distal i, és

aquest últim, el que formarà l'axonema. La peça intermèdia presenta una beina de mitocondris (de 70-100 mitocondris per espermatozoide) que envolten l'axonema i confereixen l'energia necessària per a la mobilitat espermàtica. La peça principal presenta una beina fibrosa que rodeja l'axonema. Finalment, la peça terminal de l'espermatozoide només conté l'axonema (Figura 1.3).

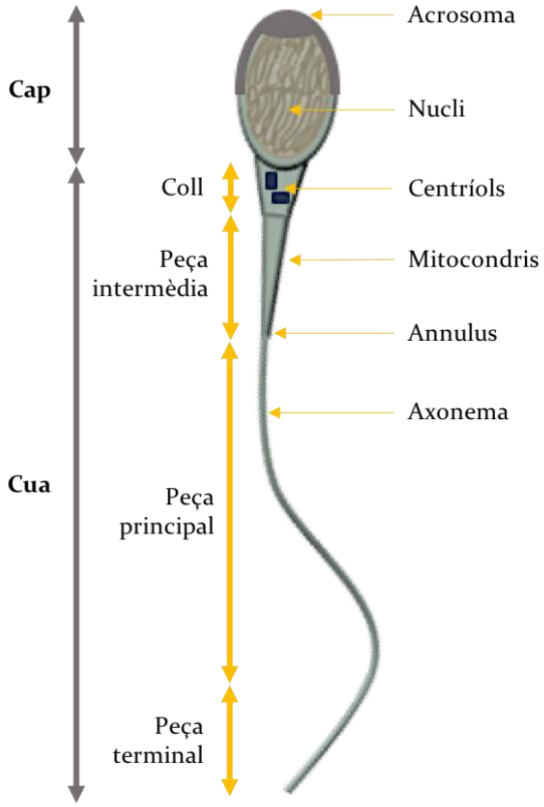


Figura 1.3. Estructures específiques de l'espermatozoide. El cap de l'espermatozoide està format pel nucli i l'acrosoma. La cua o flagel de l'espermatozoide es divideix en el coll, la peça intermèdia, la peça principal i la peça terminal. Basat en (Holstein et al., 2003).

1.1.2. Regulació hormonal de l'espermatogènesi

L'espermatogènesi està regulada hormonalment per l'eix hipotàlem-hipòfisi-gònada. L'hipotàlem segrega l'hormona alliberadora de gonadotropina (GnRH) que estimula l'alliberació, per part de la hipòfisi, de l'hormona luteïnitzant (LH) i de l'hormona estimuladora del fol·licle (FSH) que viatgen fins al testicle. Mentre que la LH té com a diana les cèl·lules de Leydig que són estimulades per a què fabriquin testosterona, la FSH estimula les cèl·lules de Sertoli per a què fabriquin inhibina B (Figura 1.4) (Holdcraft and Braun, 2004; Plant, 2015).

La testosterona alliberada per les cèl·lules de Leydig, quan assoleix els nivells requerits, inhibeix la producció tan de GnRH per part de l'hipotàlem, com de LH per part de la hipòfisi. Així mateix, la LH també té un mecanisme de regulació negativa per inhibir la producció de GnRH a l'hipotàlem. La inhibina B secretada per les cèl·lules de Sertoli, també juga un paper important en la regulació negativa, ja que té la capacitat d'inhibir la producció de FSH (Ballescà *et al.*, 2000). Tanmateix, la FSH també regula negativament la producció de GnRH (Figura 1.4) (Holdcraft and Braun, 2004; Plant, 2015).

Per a què l'espermatogènesi tingui lloc es requereixen elevats nivells de testosterona i la presència de FSH que presenten una influència reguladora positiva, creant el microambient adient. Degut a que les cèl·lules germinals no presenten receptors ni de testosterona ni FSH, són les cèl·lules de Sertoli les que intervenen en el suport físic i nutricional a les cèl·lules germinals (Holdcraft and Braun, 2004; Plant, 2015).

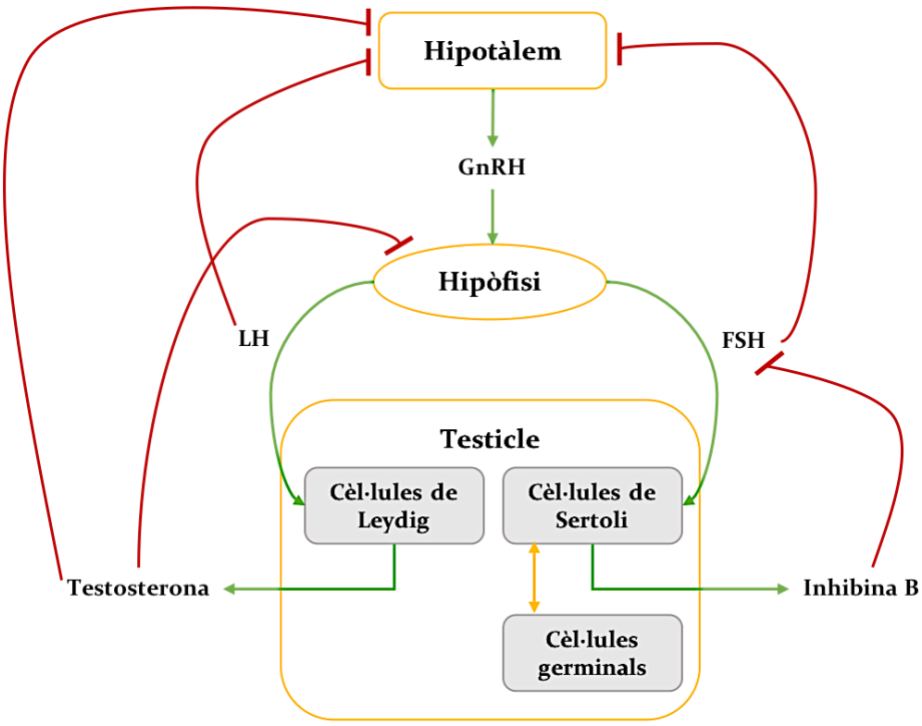


Figura 1.4. Regulació hormonal de l'espermatogènesi. L'hipotàlem secreta hormona alliberadora de gonadotropina (GnRH) i aquesta, estimula la hipòfisi per a què produeixi hormona luteïnitzant (LH) i hormona estimuladora del fol·licle (FSH). La LH té com a diana les cèl·lules de Leydig que són estimulades per a què fabriquin testosterona, mentre que la FSH estimula les cèl·lules de Sertoli per a què fabriquin inhibina B. La testosterona alliberada, quan assoleix els nivells requerits, inhibeix la producció tan de GnRH per part de l'hipotàlem, com de LH per part de la hipòfisi. Així mateix, la LH també regula negativament la producció de GnRH a l'hipotàlem. La inhibina B secretada per les cèl·lules de Sertoli, inhibeix la producció de FSH i la FSH també inhibeix la producció de GnRH. Adaptat de (Holdcraft and Braun, 2004).

1.2. Maduració, capacitació i reacció acrosòmica de l'espermatozoide

Un cop l'espermatozoide és alliberat al lumen del túbul seminífer encara és una cèl·lula immadura que li manca la motilitat i l'habilitat de fecundar l'oòcit. Per tal d'adquirir aquestes capacitats, l'espermatozoide ha de passar dos processos extra-testiculars, la maduració que té lloc a l'epidídim i al llarg del tracte sexual masculí i, la capacitació i la reacció acrosòmica que tenen lloc al tracte sexual femení (Gervasi and Visconti, 2017).

1.2.1. Maduració de l'espermatozoide en el tracte sexual masculí

Un cop els espermatozoides són produïts al testicle, aquests s'emmagatzemen fins a dues setmanes a l'epidídim (Gadea *et al.*, 2013; Breton *et al.*, 2016). L'epidídim és un òrgan tubular que es troba situat i connectat a la part superior dels testicles i és on té lloc principalment la maduració de l'espermatozoide. L'epidídim es pot dividir en tres parts principals: la porció *caput* que és la part proximal més propera al testicle, la porció *corpus* que és la regió intermèdia i la porció *cauda* que és la regió més distal que es troba connectada al *vas deferens,* on s'emmagatzemaran els espermatozoides fins al moment de l'ejaculació (Lin and Troyer, 2014). L'epiteli de l'epidídim absorbeix la major part del volum de líquid secretat prèviament per la *rete testis* i provoca l'augment de la concentració de 10^6 espermatozoides / ml a la *rete testis* a 10^9 espermatozoides / ml a la porció *cauda* de l'epidídim (Dacheux and Dacheux, 2014; Dacheux *et al.*, 2016).

Degut a l'alta segmentació de l'epidídim, el líquid epididimari alliberat al lumen per les cèl·lules epitelials varia entre regions, formant un microambient diferenciat en cada un dels segments que permetrà la maduració progressiva de l'espermatozoide (Figura 1.5) (Bedford, 2004; Sullivan and Mieusset, 2016). Així doncs, els espermatozoides pateixen un procés de remodelació on perden i modifiquen lípids, RNAs i proteïnes i n'adquireixen d'altres a través de l'intercanvi d'informació amb l'epidídim i els seus segments (Figura 1.5), bé per l'absorció per part de l'espermatozoide de proteïnes solubles, per la interacció de la cèl·lula espermàtica amb vesícules extracel·lulars alliberades per l'epiteli epididimari (epididimosomes) i/o pel contacte directe dels espermatozoides amb l'epitelial de l'epidídim (Cornwall, 2009; Sullivan and Saez, 2013; Sullivan and Mieusset, 2016; Gervasi and Visconti, 2017; Sullivan and Belleannée, 2017; Zhou *et al.*, 2018; Nixon *et al.*, 2019; James *et al.*, 2020).

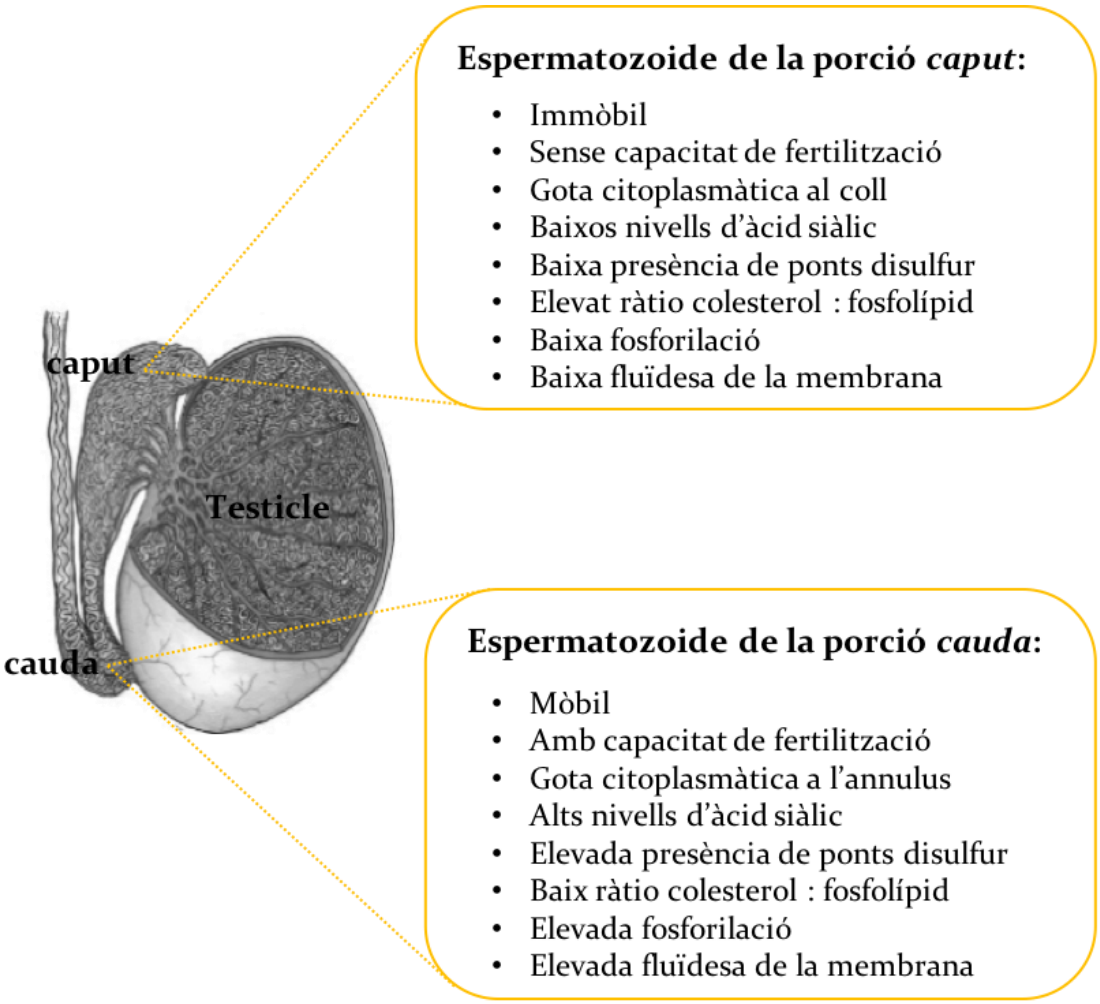


Figura 1.5. Maduració de l'espermatozoide a l'epidídim. Característiques diferencials dels espermatozoides a l'inici de la maduració a la porció inicial *caput* i a la porció final *cauda*. Adaptat de (Gervasi and Visconti, 2017).

Durant la maduració, algunes de les proteïnes contingudes prèviament pateixen modificacions posttraduccionals (PTMs) (Figura 1.5). Un exemple és l'oxidació dels grups tiol formant ponts disulfur que es dona principalment a proteïnes de la cua de l'espermatozoide que estabilitzen el flagel i a les protamines que empaqueten fortament i estabilitzen la cromatina del nucli (Veure apartat 1.4.3; veure annex 1) (Oliva, 2006; Gervasi and Visconti, 2017). També, la fosforilació de certes proteïnes podria tenir un rol rellevant en l'adquisició de l'habilitat de l'espermatozoide de fertilitzar l'oòcit. Per exemple, l'àmplia fosforilació al llarg del trànsit de l'epidídim de la proteïna *izumo Sperm-Egg Fusion 1* (IZUMO1) que és essencial per la fusió de l'espermatozoide amb l'oòcit (Inoue *et al.*, 2005; Baker *et al.*, 2012). Durant aquest procés de maduració posttesticular, també s'adquireixen noves proteïnes i RNAs, entre elles proteïnes essencials per la fertilització com són proteïnes d'unió a la zona pel·lúcida, la proteïna d'unió acrosina, la d*isintegrin and metalloproteinase domain-containing protein 7* (ADAM7) i la *cysteine-rich secretory protein 1* (CRISP1) (Oh *et*

al., 2009; Cohen *et al.*, 2011; Maldera *et al.*, 2014; Sullivan and Mieusset, 2016; Barrachina *et al.*, 2019; Nixon *et al.*, 2019; James *et al.*, 2020). A part de l'adquisició de noves proteïnes, RNAs i la incorporació de PTMs, la gota citoplasmàtica de l'espermatozoide migra de la regió del coll a l'*annulus* (Gervasi and Visconti, 2017). A més, la membrana de l'espermatozoide pateix diferents canvis moleculars (Figura 1.5) com ara l'augment de la concentració intracel·lular d'àcid siàlic i la disminució del ràtio colesterol : fosfolípids que n'incrementa la fluïdesa (Gervasi and Visconti, 2017).

Per tant, els fenòmens que tenen lloc durant la maduració de l'espermatozoide són principalment: (i) l'adquisició de la mobilitat espermàtica; (ii) l'adquisició de la capacitat de l'espermatozoide de realitzar la reacció acrosòmica, reconèixer i unir-se a la zona pel·lúcida i fusionar-se amb l'oòcit; i (iii) canvis estructurals, morfològics i de contingut.

En el moment de l'ejaculació, els espermatozoides són alliberats des de l'epidídim al conducte deferent on s'hi incorporaran les secrecions provinents de les anomenades glàndules sexuals accessòries formades per la pròstata (≈ 25 – 30 % del volum del plasma seminal), les vesícules seminals (≈ 65 – 75 % del volum del plasma seminal) i les glàndules bulbouretrals i periuretrals (< 1 % del volum del plasma seminal). Tot i que clàssicament s'havia assumit que el plasma seminal era solament un mitjà de transport i nutrició per a l'espermatozoide (Suarez and Wolfner, 2017), s'ha demostrat que la seva interacció és crucial per a la maduració espermàtica i l'adquisició de la mobilitat i el potencial de fertilització (Sullivan and Mieusset, 2016; Barrachina *et al.*, 2019; Björkgren and Sipilä, 2019).

1.2.2. Capacitació i reacció acrosòmica en el tracte sexual femení

La capacitació dels espermatozoides té lloc al tracte sexual femení una vegada els espermatozoides són ejaculats. Primerament, milions d'espermatozoides travessen el conducte vaginal fins a arribar a l'úter (Figura 1.6A). A la cavitat uterina, els espermatozoides perden els factors decapacitants de la superfície adquirits durant la maduració. N'és un exemple la pèrdua de colesterol de la membrana plasmàtica per part de molècules acceptores com l'albúmina i la presència d' espècies reactives d'oxigen (ROS) del tracte sexual femení. Aquesta pèrdua provoca l'augment de la fluïdesa de la membrana i l'augment a la permeabilitat pel ió Ca^{2+} (Aitken and Nixon, 2013; de Jonge, 2017; Molina *et al.*, 2018). Seguidament es dipositen a l'oviducte on s'uneixen a les cèl·lules epitelials (Figura 1.6A) i mitjançant una senyal endocrina, segurament progesterona, els espermatozoides adquireixen un canvi bioquímic caracteritzat per un increment dels nivells de monofosfat d'adenosina cíclic (cAMP), un increment de la generació de ROS i un increment massiu de la fosforilació de

residus tirosina per part de diferents quinases. Com a conseqüència d'aquests canvis, s'obren els canals de Ca^{2+} anomenats canals de cations de l'espermatozoide (CatSper) i hi ha una alliberació de Ca^{2+} intracel·lular que provoca l'adquisició de la hipermotilitat (Figura 1.6A) (Aitken and Nixon, 2013; de Jonge, 2017; Molina *et al.*, 2018; Castillo *et al.*, 2019). S'han descrit vàries proteïnes implicades en l'adquisició de la hiperactivació com són els canals de Ca^{2+} específics d'espermatozoides CatSpers esmentats anteriorment, les *A-kinase anchor* 3 i 4 (AKAP3 i AKAP4) o la *protein kinase A* (PKA) (Ficarro *et al.*, 2003; Singh and Rajender, 2015; Williams *et al.*, 2015; Castillo *et al.*, 2018, 2019; Dey *et al.*, 2019). En ratolí, el *knock-out* (KO) de la *cation channel sperm-associated protein 1* (*Catsper1)* és incapaç de desenganxar-se de l'epiteli de l'oviducte degut a que no adquireix l'estat d'hipermotilitat (Ho *et al.*, 2009) i, en humà, s'ha determinat que mutacions al gen de *CATSPER1* provoca infertilitat (Avenarius *et al.*, 2009). La hiperactivació de la cua de l'espermatozoide, on els moviments passen del batec simètric de baixa amplitud al batec ràpid asimètric de gran amplitud (Yanagimachi *et al.*, 1994), permet l'avançament de l'espermatozoide de l'epiteli oviductal a la regió on hi ha l'oòcit i, segurament, ajuda a la penetració de la zona pel·lúcida (Figura 1.6A) (Aitken and Nixon, 2013). Finalment, un cop els espermatozoides arriben a l'oòcit, aquests realitzen la reacció acrosòmica que permet la penetració de la zona pel·lúcida (Figura 1.6B) i la fertilització (de Jonge, 2017).

La reacció acrosòmica és el procés d'exocitosi de la vesícula acrosòmica que té lloc quan l'espermatozoide s'aproxima a l'oòcit, en resposta de factors solubles de la dona com ara la progesterona i a la presència de les glicoproteïnes de membrana de la zona pel·lúcida de l'oòcit. Aquesta interacció desencadena un augment de Ca^{2+} intracel·lular que provoca la reacció acrosòmica. La membrana exterior de la vesícula acrosòmica es fusiona amb la membrana plasmàtica de l'espermatozoide i s'allibera tot el contingut de la vesícula, principalment enzims hidrolítics i proteolítics com ara l'acrosina i la hialuronidasa (Figura 1.6B). La membrana acrosòmica interna quedarà exposada amb receptors de membrana que permetran la penetració de l'espermatozoide a l'oòcit (Figura 1.6B) i per tant, la fusió de les membranes que culminarà amb la fertilització (Brucker and Lipford, 1995; Patrat *et al.*, 2000; Chakravarty *et al.*, 2008; Ganguly *et al.*, 2010; Wu and Sampson, 2014; Castillo *et al.*, 2019).

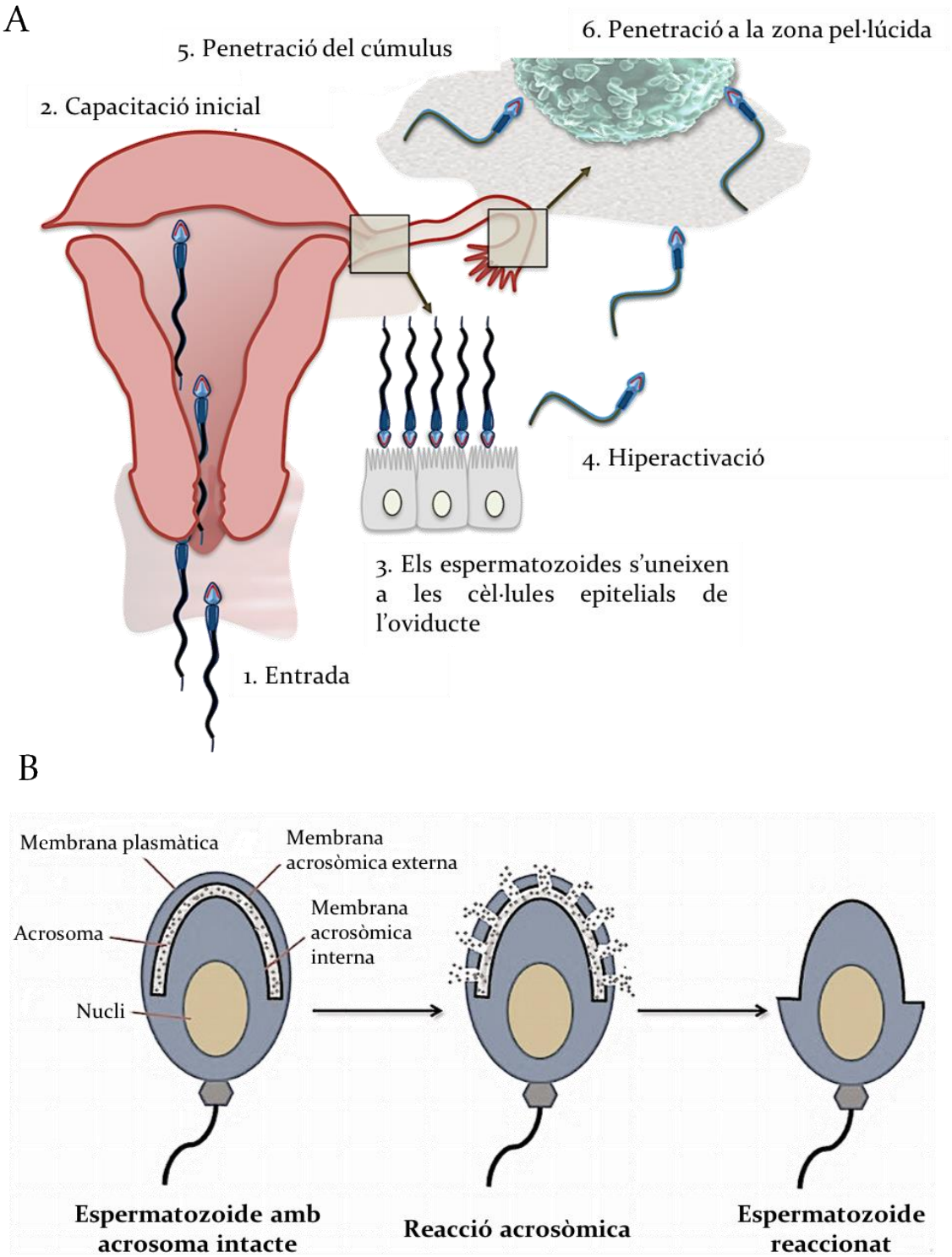


Figura 1.6. Capacitació i reacció acrosòmica de l'espermatozoide. A) Els espermatozoides entren al tracte genital femení. A la cavitat uterina realitzen la capacitació inicial. Seguidament es dipositen a les cèl·lules epitelials de l'oviducte on adquireixen la hiperactivació. Un cop hiperactivats, els espermatozoides penetren les cèl·lules del cúmulus i realitzen la reacció acrosòmica que permet la penetració a la zona pel·lúcida i posterior fertilització. Adaptat de (Aitken and Nixon, 2013). B) En la reacció acrosòmica, la membra exterior de la vesícula acrosòmica es fusiona amb la membrana plasmàtica de l'espermatozoide i s'allibera tot el contingut de la vesícula i, finalment queda exposada la membrana acrosòmica interna. Adaptat de (Wu and Sampson, 2014).

1.3. Fertilització i desenvolupament embrionari primerenc

Un cop l'espermatozoide duu a terme la capacitació i la reacció acrosòmica, culmina en el procés de fertilització, en el qual l'espermatozoide i l'oòcit, ambdues cèl·lules haploides, es combinen per generar el zigot totipotent diploide (Figura 1.7). En aquest moment, començarà el desenvolupament embrionari primerenc que acabarà amb la implantació de l'embrió a la paret de l'úter.

1.3.1. Fertilització

Primerament l'espermatozoide penetra el cúmulus i entra en contacte amb la zona pel·lúcida de l'oòcit. Específicament, les proteïnes glicosilades de la zona pel·lúcida *zona pellucida sperm-binding protein* (ZP1, ZP2, ZP3 i ZP4), sobretot ZP3, interactuen amb proteïnes de l'espermatozoide com l'heterodímer fertilina (composat pels membres de la família *disintegrin and metalloproteinase domain-containing protein* 1B i 2; ADAM1B/ADAM2), la *zona pellucida-binding protein 2* (ZPBP2), la *disintegrin and metalloproteinase domain-containing protein* 3 (ADAM3) i la *zona pellucida sperm-binding protein 3 receptor* (ZP3R), entre altres (Torabi *et al.*, 2017; Satouh and Ikawa, 2018). Aquesta interacció permet que es dugui a terme la reacció acrosòmica i gràcies als enzims amb activitat hialuronidasa que s'expulsen, com ara l'acrosina, l'espermatozoide aconsegueix travessar la zona pel·lúcida.

Un cop travessada la zona pel·lúcida, l'espermatozoide arriba a l'espai perivitel·lí. Allà s'unirà al microvil·li de la membrana de l'oòcit que permet que hi hagi una disminució de la repulsió entre les membranes plasmàtiques (Satouh and Ikawa, 2018). A la membrana de l'oòcit hi ha clústers de les proteïnes de membrana *CD9 antigen* (CD9) que es troben repartits al llarg de tot l'oòcit que permeten una millor adhesió de l'espermatozoide per la regió equatorial (Figura 1.7). Gràcies a la reacció acrosòmica, proteïnes específiques de l'espermatozoide queden exposades, de les quals, la proteïna espermàtica IZUMO1 interactua amb la proteïna de l'oòcit *sperm-egg fusion protein Juno* (JUNO; Figura 1.7). Un cop s'han unit IZUMO1 i JUNO, s'acumula CD9 a la regió de contacte i propicia la formació d'un clúster de proteïnes de membrana que participen en la fusió. Finalment, les membranes d'ambdós gàmetes es fusionen (Figura 1.7) (Inoue *et al.*, 2005; Bianchi *et al.*, 2014; Aydin *et al.*, 2016; Ohto *et al.*, 2016; Trebichalská and Holubcová, 2020). S'ha observat que els ratolins *Izumo1*$^{-/-}$ són completament infèrtils, ja que els seus espermatozoides poden penetrar la zona pel·lúcida però no fusionar-se amb l'oòcit (Inoue *et al.*, 2005). Recentment s'ha proposat, en ratolí, que les proteïnes *WD repeats and SOF1 domain containing* (Sof1), *transmembrane protein 95* (Tmem95) i *sperm acrosome membrane-associated protein 6* (Spaca6) poden funcionar cooperativament amb

Izumo1, juntament amb altres molècules en el procés de fusió de l'espermatozoide i l'oòcit (Barbaux *et al.*, 2020; Noda *et al.*, 2020).

Un cop l'espermatozoide ha entrat a l'oòcit, la proteïna *phospholipase C zeta* (PLCζ) que és específica del cap de l'espermatozoide, és alliberada dins l'oòcit i indueix les anomenades oscil·lacions de Ca^{2+} que es basen en l'augment i disminució cíclic de Ca^{2+} intracel·lular post-fertilització i, això, és el detonant del començament del desenvolupament embrionari primerenc (Figura 1.7) (Ramadan *et al.*, 2012). La família *phospholipase C* (PLC) digereix el fosfatidil-inositol-4,5-bifosfat (PIP_2) en inositol trifosfat (IP_3) i diacilglicerol en presència de Ca^{2+} i, l'augment de IP_3 desencadena les oscil·lacions de Ca^{2+}. La isoforma PLCζ és l'única forma que s'activa amb uns nivells de Ca^{2+} tan baixos com els presents en l'oòcit (Satouh and Ikawa, 2018).

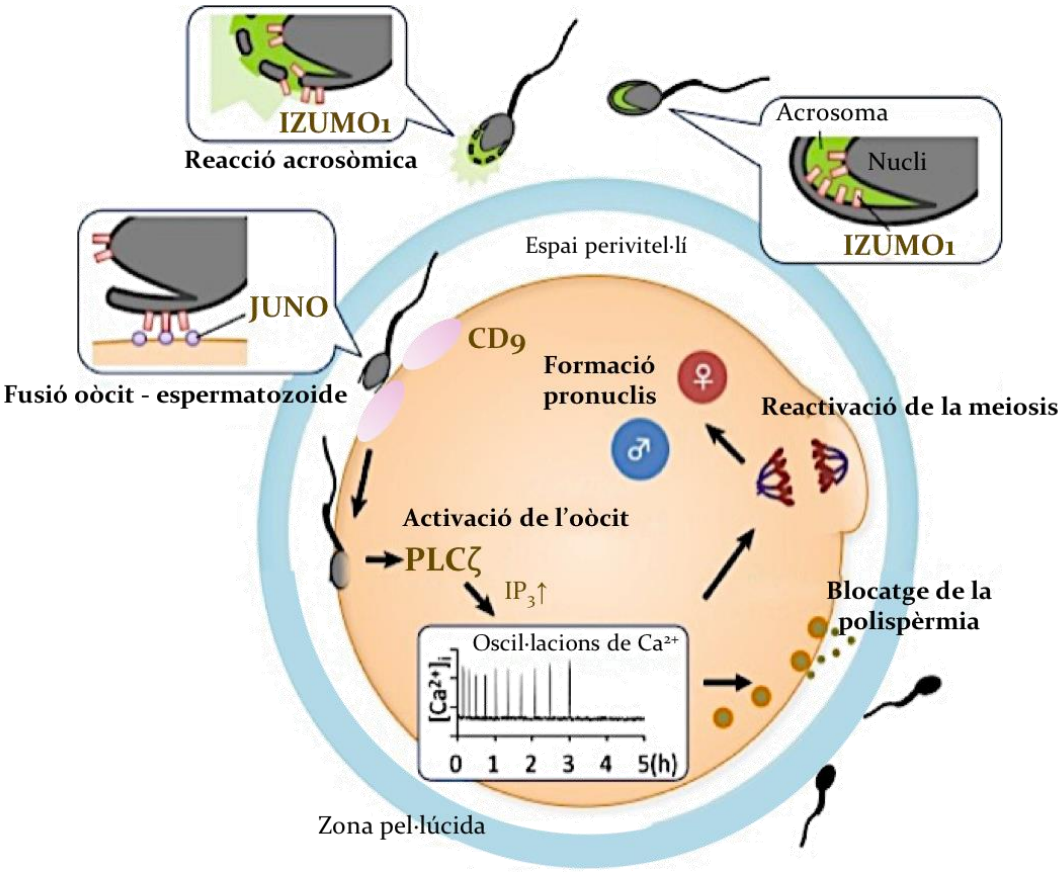


Figura 1.7. Procés de fertilització. Després de la reacció acrosòmica i un cop l'espermatozoide travessa la zona pel·lúcida, la proteïna *izumo sperm-egg fusion 1* (IZUMO1) és exterioritzada a la membrana de l'espermatozoide i interactua amb les proteïnes *sperm-egg fusion protein Juno* (JUNO) i *CD9 antigen* (CD9), presents a la membrana de l'oòcit, fet que permet la fusió de les membranes. Un cop les membranes es fusionen, l'espermatozoide allibera *phospholipase C zeta* (PLCζ) que provoca un augment de inositol trifosfat (IP_3) en l'oòcit que desencadena les oscil·lacions de Ca^{2+}. Aquestes oscil·lacions permetran la reactivació de la meiosis en l'oòcit i posterior formació dels dos pronuclis i el blocatge de la polispèrmia. Adaptat de (Satouh and Ikawa, 2018).

Com s'ha mencionat, PLCζ desencadena les oscil·lacions de Ca^{2+} i, aquestes, dirigeixen una sèrie d'esdeveniments seqüencials com són, entre altres, el blocatge de la polispèrmia i la represa de la meiosi de l'oòcit que es trobava aturada en metafase II (Figura 1.7) (Trebichalská and Holubcová, 2020). El blocatge de la polispèrmia es caracteritza, primerament, per una despolarització transitòria de la membrana de l'oòcit fertilitzat que passa de ser negativa a ser positiva (Figura 1.8), tot i que en humans no està del tot clar si té lloc (Trebichalská and Holubcová, 2020; Wozniak and Carlson, 2020). També, un cop IZUMO1 i JUNO s'han unit, les molècules de JUNO de la membrana de l'oòcit són eliminades per exocitosi a través de vesícules extracel·lulars (Figura 1.8), la qual cosa evita que més espermatozoides es puguin unir a la membrana de l'oòcit (Chalbi *et al.*, 2014). Les oscil·lacions de Ca^{2+} també provoquen l'exocitosi dels grànuls corticals a l'espai perivitel·lí (Figura 1.8). El contingut alliberat dels grànuls corticals conté proteases que trenquen la glicoproteïna ZP2 i provoca que la zona pel·lúcida es converteixi en impermeable (Figura 1.8), evitant l'entrada d'espermatozoides i esdevenint una capa protectora pel futur embrió (Liu, 2011). Finalment, l'últim mecanisme de blocatge de la polispèrmia és l'alliberació de Zn^{2+} per exocitosis (Figura 1.8). Aquesta disminució del Zn^{2+} intracel·lular, permet la reactivació de la meiosis. També, juntament amb les oscil·lacions de Ca^{2+}, el Zn^{2+} que s'acumula a l'exterior també contribueix als canvis que es produeixen a la zona pel·lúcida (Duncan *et al.*, 2016).

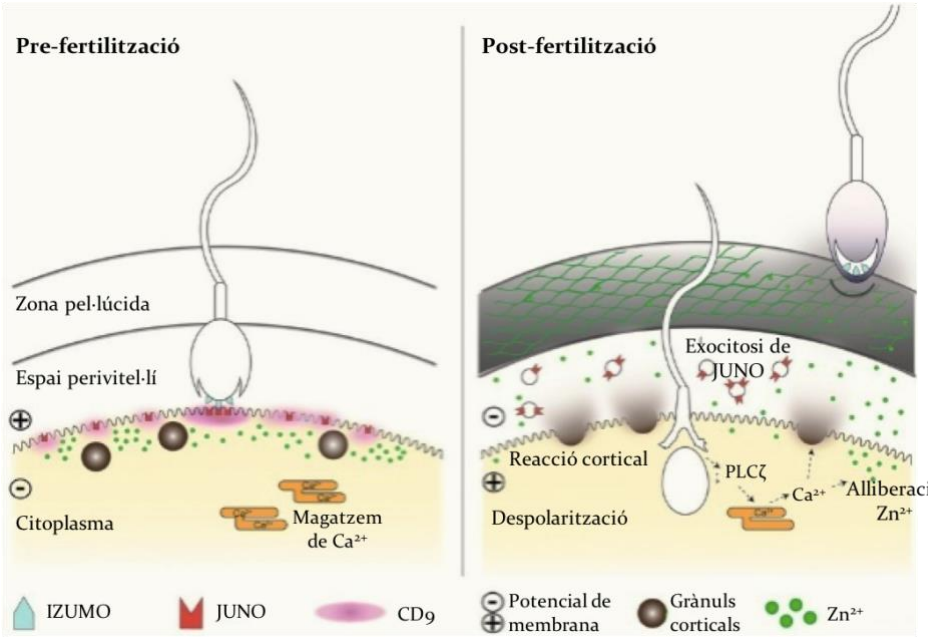


Figura 1.8. Blocatge de la polispèrmia. L'entrada de l'espermatozoide i l'alliberació de *phospholipase C zeta* (PLCζ) desencadena les oscil·lacions de Ca^{2+} que provoquen la despolarització transitòria de la membrana de l'oòcit, l'exocitosi de JUNO i la reacció cortical i l'alliberació de Zn^{2+}. Tots aquests canvis provoquen que la zona pel·lúcida es torni impermeable a l'entrada d'altres espermatozoides. IZUMO1: i*zumo sperm-egg Fusion 1*; JUNO: *sperm-egg fusion protein Juno*; CD9: *CD9 antigen*. Adaptat de (Trebichalská and Holubcová, 2020).

1.3.2. Desenvolupament embrionari primerenc

El desenvolupament embrionari primerenc comença amb la formació dels dos pronuclis i acaba amb la implantació de l'embrió a l'endometri matern. Té lloc durant el trànsit de l'embrió per l'oviducte i culmina amb l'entrada a l'úter i posterior implantació a l'endometri. En humans aquest procés té una durada aproximada de 7 dies. (Figura 1.9).

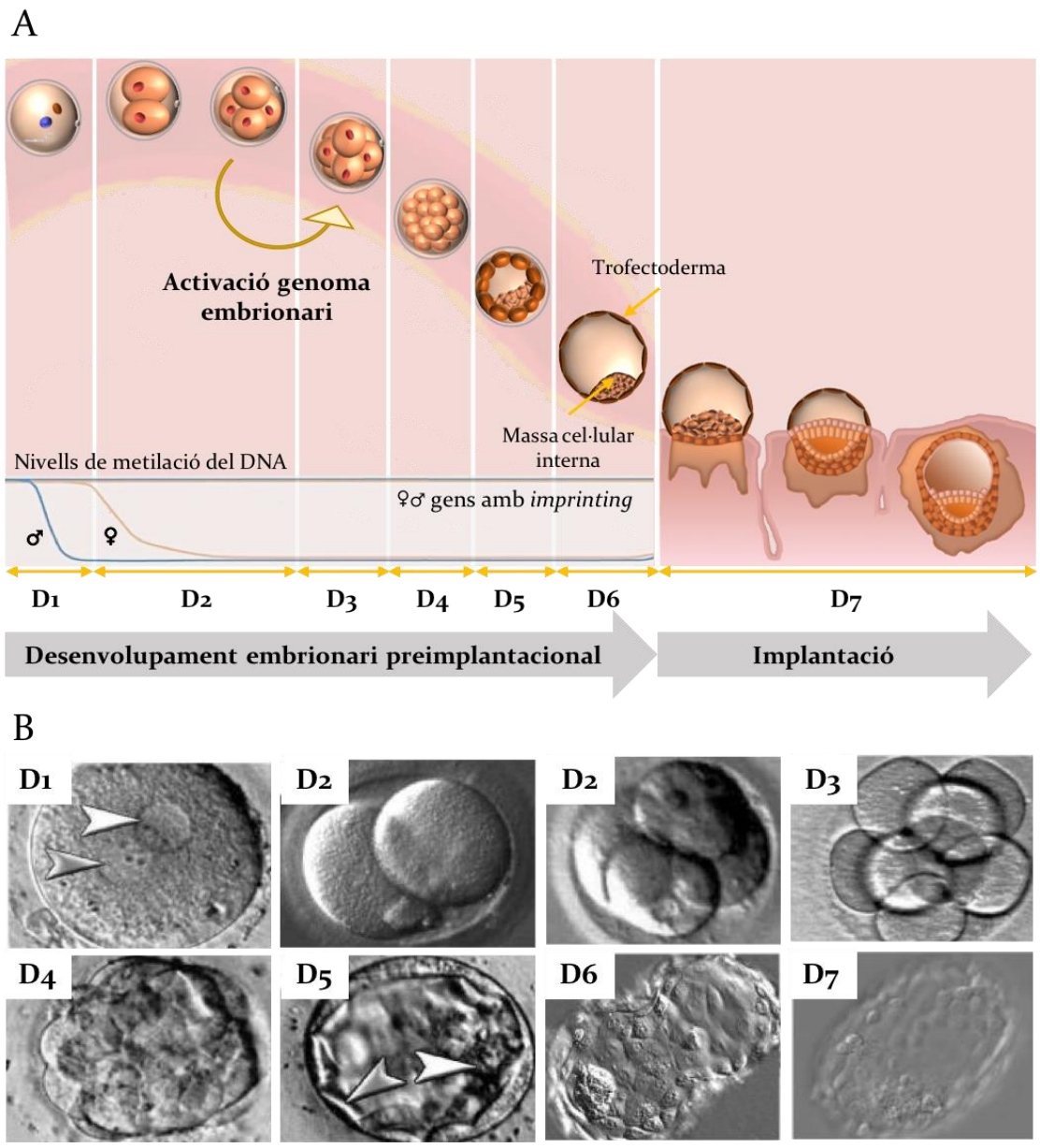


Figura 1.9. Desenvolupament embrionari primerenc. A) Esquematització del desenvolupament embrionari preimplantacional i la implantació des de la fertilització fins al dia 7 de desenvolupament embrionari i, patró dels nivells de metilació dels genomes patern i matern. Figura cedida en part per Alberto de la Iglesia i adaptat de (Castillo *et al.*, 2018). B) Fotografies reals del desenvolupament embrionari primerenc des del dia 1 fins al dia 7 (D1-D7). Les fletxes en D1 marquen els dos pronuclis. La fletxa grisa en D5 marca el trofectoderma i la fletxa blanca en D5 marca la massa cel·lular interna. Adaptat de (Niakan *et al.*, 2012).

Un cop l'oòcit és fertilitzat per l'espermatozoide, primerament es formen els dos pronuclis i es genera l'anomenat zigot. Aquest procés té lloc al cap d'unes 16 h post-fertilització. L'espermatozoide perd la membrana plasmàtica, però el seu DNA es troba majoritàriament empaquetat amb les protamines (Veure apartat 1.4.3.2). L'ambient reductor de l'oòcit, contribueix a que les protamines es desempaquetin del DNA patern i siguin substituïdes per histones de l'oòcit (Oliva and Dixon, 1991; Oliva, 2006; Jones *et al.*, 2012). Finalment, el DNA patern s'envolta d'una nova membrana nuclear i queda el pronucli patern format. Paral·lelament i degut a les oscil·lacions de Ca^{2+} que es donen en la fertilització (Veure apartat 1.3.1), l'oòcit, que es trobava aturat en metafase II, reactiva la meiosis que culmina amb l'extrusió del segon corpuscle polar i permet la formació del pronucli matern. Un cop els pronuclis han estat formats, aquests migren fins al centre de l'oòcit gràcies a la presència de microfilaments que els acompanyen (Bustin, 2015). En humans, els dos pronuclis no es fusionen sinó que realitzen l'anomenada sincronització metafàsica on els dos genomes s'encaren i formen el fus mitòtic i comencen així les primeres divisions embrionàries.

En l'estadi de zigot hi ha una reprogramació epigenètica extensiva que permetrà que l'embrió en desenvolupament sigui totipotent (Reik and Walter, 2001; Reik *et al.*, 2001). El DNA del pronucli masculí començarà a patir una desmetilació activa, mentre que la del DNA del pronucli femení serà passiva, mantenint les marques de *l'imprinting* parental i les seqüències repetitives (Figura 1.9) (Wei *et al.*, 2014). En el pronucli masculí, les regions amb *imprinting* patern es trobaran metilades, mentre que les regions amb *imprinting* matern es trobaran desmetilades. El contrari succeeix amb el pronucli femení, on les regions amb *imprinting* matern es trobaran metilades i les regions amb *imprinting* patern es trobaran desmetilades. Aquesta metilació selectiva permetrà l'expressió específica de gens d'origen patern o matern durant els primers estadis de l'embriogènesi. Aquesta ràpida desmetilació del genoma patern pot ser que permeti la transcripció primerenca observada en l'estadi de 2 cèl·lules, previ a l'activació del genoma embrionari (Santos and Dean, 2004; Vassena *et al.*, 2011).

La primera divisió mitòtica del zigot, que passarà a tenir 2 cèl·lules o blastòmers, ocorre el dia 2 del desenvolupament embrionari. Seguidament i durant el mateix dia, es tornarà a dividir passant a tenir 4 blastòmers i l'endemà es dividirà a 8 blastòmers (Figura 1.9). En humà, és entre l'estadi de 4 – 8 cèl·lules en què té lloc l'activació del genoma embrionari humà (Figura 1.9), mentre que en ratolí succeeix molt més primerencament, a l'estadi de 2 cèl·lules, entre 26 – 29 h post-fertilització (Niakan *et al.*, 2012). Just abans de l'activació del genoma embrionari hi ha una eliminació massiva dels RNAs i proteïnes maternes per tal de donar pas a la maquinària embrionària, ja que fins a l'activació del genoma embrionari s'utilitza la maquinària materna (Walser and Lipshitz, 2011).

Un cop el genoma embrionari s'activa, té lloc el pas de l'estadi de 8 cèl·lules a l'estadi de mòrula en el dia 4 del desenvolupament embrionari. És en aquest estadi que es dona el fenomen d'adhesió cel·lular anomenat compactació, on els blastòmers ja no es poden distingir entre ells i les cèl·lules queden agrupades (Figura 1.9) (Niakan *et al.*, 2012). Durant el cinquè dia del desenvolupament embrionari primerenc, es forma el blastocel, que és una cavitat plena de líquid que va creixent dins l'embrió i, el blastocist queda totalment format el dia 6 del desenvolupament embrionari (Figura 1.9). En l'estadi de blastocist, les cèl·lules ja s'han diferenciat en principalment dos tipus, les cèl·lules que donaran lloc a l'embrió en si, anomenada massa cel·lular interna i les cèl·lules que donaran lloc a la placenta i les membranes amniòtiques, anomenada trofectoderma (Figura 1.9) (Niakan *et al.*, 2012). En el sisè dia del desenvolupament embrionari també té lloc l'eclosió del blastocist, el qual s'expandeix totalment i la zona pel·lúcida que l'ha envoltat durant tot el desenvolupament, s'aprima. Gràcies a la secreció d'enzims lítics i moviments d'expansió del blastocist, aquest fa un forat a la zona pel·lúcida i per contracció i expansió se n'allibera (Figura 1.9) (Shafei *et al.*, 2017).

En el dia 7 de desenvolupament embrionari, el blastocist eclosionat rau orientat d'una manera concreta vers l'úter, presenta una elevada capacitat d'adhesió i interactua amb l'epiteli de l'endometri. Aquesta interacció dispara la reacció de la decídua, que bàsicament implica una angiogènesi local de l'úter allà on el blastocist s'està unint i es comença a formar la part materna de la placenta (Diedrich *et al.*, 2007). En paral·lel, s'inhibeix el sistema immunitari matern per a què la implantació sigui satisfactòria. Un cop adherit, el blastocist primer envaeix la capa epitelial de l'endometri i finalment envaeix l'estroma (Figura 1.9). Aquest procés però, només pot tenir lloc en el període finestra en què l'endometri es troba receptiu al blastocist, que té lloc al voltant dels 6 – 8 dies després de la fertilització (Bergh and Navot, 1992; Diedrich *et al.*, 2007). Així doncs, el procés d'implantació del blastocist es divideix en tres passos seqüencials: (i) L'aposició que implica la correcta orientació del blastocist a l'endometri; (ii) L'adhesió a l'epiteli de l'endometri; i (iii) La invasió del blastocist a l'estroma endometrial (Diedrich *et al.*, 2007).

1.4. Contribució de l'espermatozoide madur a l'embrió primerenc

Tot i que tradicionalment sempre s'havia pensat que l'única funció de l'espermatozoide era el transport del genoma haploide patern fins a l'oòcit madur, recentment s'ha observat que la cèl·lula espermàtica també aporta durant la fertilització altres components, com un estímul d'activació de l'oòcit, el centríol que permetrà la primera divisió embrionària, el DNA i marques epigenètiques específiques, RNAs i proteïnes, els quals podrien ser claus en el desenvolupament embrionari primerenc (Jodar *et al.*, 2013, 2015, 2016a; Castillo *et al.*, 2015, 2018; Jodar, 2019). Així doncs, tot i la concepció clàssica de que només s'utilitzen les proteïnes i RNAs materns fins a l'activació del genoma embrionari, les proteïnes i RNAs aportats inicialment per l'espermatozoide també podrien jugar un paper important, ja sigui per la contribució de proteïnes, per la traducció d'RNAs intactes utilitzant la maquinària traduccional materna que es troba activa o per la regulació de l'expressió gènica mitjançant RNAs no codificants, proteïnes o PTMs d'aquestes (Miller, 2015; Ntostis *et al.*, 2017; Castillo *et al.*, 2018; Jodar, 2019).

1.4.1. Estímul d'activació de l'oòcit

Fins al moment de la fertilització, l'oòcit roman en un estadi quiescent aturat en l'estadi de metafase II de la meiosi. Un cop l'espermatozoide entra a l'oòcit i es produeix la fertilització (Veure apartat 1.3.1), s'indueixen les oscil·lacions de Ca^{2+} que permeten la finalització de la meiosi oocitària que culmina amb l'extrusió del segon corpuscle polar i la formació del pronucli femení (Satouh and Ikawa, 2018; Trebichalská and Holubcová, 2020). La inducció d'aquestes oscil·lacions de Ca^{2+} ve donada per un factor que aporta exclusivament la cèl·lula espermàtica en el moment de la fertilització (Veure apartat 1.3.1). Actualment se sap que l'estímul d'activació de l'oòcit és la proteïna de l'espermatozoide PLCζ (Yoda *et al.*, 2004; Yoon *et al.*, 2008; Heytens *et al.*, 2009; Kashir *et al.*, 2012; Ramadan *et al.*, 2012; Hachem *et al.*, 2017; Nozawa *et al.*, 2018). En ratolí, s'ha observat que la injecció del transcrit de *phospholipase C zeta 1* (*Plcz1*) és traduït i provoca l'activació de l'oòcit *in vitro* (Yoda *et al.*, 2004). De la mateixa manera, s'ha descrit que una alteració en l'expressió de PLCζ provoca una reducció o absència de les oscil·lacions de Ca^{2+} (Yoon *et al.*, 2008; Heytens *et al.*, 2009; Kashir *et al.*, 2012; Ramadan *et al.*, 2012). De fet, en ratolins *Plcz1*$^{-/-}$, l'espermatogènesi, la motilitat espermàtica, la reacció acrosòmica i la fusió oòcit – espermatozoide són similars als ratolins *wild-type* (WT), però hi ha una fallida en la inducció de les oscil·lacions de Ca^{2+} després de l'ús de fecundació *in vitro* amb injecció intracitoplasmàtica (FIV-ICSI), tot i que mitjançant concepció natural o fecundació *in vitro* convencional (FIV convencional), en alguns casos es pot arribar

a activar l'oòcit, però amb una elevada taxa de polispèrmia (Hachem *et al.*, 2017; Nozawa *et al.*, 2018). En humans s'han descrit múltiples mutacions en el gen de *PLCζ* en individus que presenten fallida de fertilització després de l'ús de FIV-ICSI, en alguns dels quals es rescatava el procés si s'aplicava l'activació oocitària assistida (Torra-Massana *et al.*, 2019). Tot i que PLCζ presenta un rol clau en l'activació de les oscil·lacions de Ca^{2+} i en l'activació de l'oòcit, alguns ratolins *Plcz1*$^{-/-}$ encara són capaços d'obtenir descendència (Nozawa *et al.*, 2018), la qual cosa indica que hi ha altres factors de l'espermatozoide que poden compensar l'activitat de PLCζ. Aquesta hipòtesi es troba recolzada en què, en humans, hi ha pacients amb fallides de fertilització que no presenten alteracions en el gen de *PLCζ*, ni alteracions en la seva localització o contingut a nivell proteic (Ferrer-Vaquer *et al.*, 2016; Torra-Massana *et al.*, 2019).

1.4.2. Centríol

De manera general, el centrosoma es troba composat de proteïnes associades i un parell de centríols que formen el centre organitzador de microtúbuls de la cèl·lula que és l'encarregat de la correcta segregació dels cromosomes i la formació del fus mitòtic durant la divisió cel·lular. Durant les etapes finals de l'espermatogènesi, el centrosoma es divideix en dos centríols, el proximal i el distal. El centríol distal, juntament amb proteïnes associades al centrosoma és eliminat durant l'espermatogènesi i forma l'axonema (Veure apartat 1.1.1.4). Per tant, l'espermatozoide madur només conserva el centríol proximal (Schatten, 1994; Manandhar *et al.*, 2005). L'oòcit, per la seva part, durant l'oogènesi perd els dos centríols, tot i que manté certes proteïnes associades al centrosoma (Schatten, 1994; Manandhar *et al.*, 2005; Inoue *et al.*, 2018). Un cop l'espermatozoide fertilitza l'oòcit i es formen els dos pronuclis, la hipòtesi més acceptada és que el centríol proximal, que ha aportat la cèl·lula espermàtica, es divideix, es duplica i, amb l'ajuda de les proteïnes del centrosoma de l'oòcit forma un centrosoma funcional per l'embrió que s'organitza per formar el fus mitòtic i permet la divisió cel·lular (Schatten, 1994; Manandhar *et al.*, 2005; Niakan *et al.*, 2012; Bustin, 2015; Inoue *et al.*, 2018). Recentment, però, s'ha identificat que el centríol distal de l'espermatozoide és capaç de remodelar-se i posicionar-se, juntament amb el centríol proximal i les proteïnes associades de l'oòcit i formar el centrosoma funcional del zigot (Fishman *et al.*, 2018; Avidor-Reiss and Fishman, 2019). Tot i així, són necessaris més estudis per tal de validar quin és el mecanisme exacte de la formació del centrosoma funcional del zigot.

1.4.3. DNA i marques epigenètiques específiques

El genoma patern i matern no són equivalents i és necessària la seva contribució diferencial pel correcte desenvolupament embrionari. La modificació epigenètica més estudiada és la metilació del DNA, però n'hi ha d'altres com l'estructura única de la cromatina espermàtica que és necessària per un correcte desenvolupament embrionari (Rousseaux *et al.*, 2005; Oliva, 2006; Auclair and Weber, 2012; Oliva and Ballescà, 2012; Castillo *et al.*, 2014a; Jodar and Oliva, 2014; Carrell *et al.*, 2016).

1.4.3.1. Metilació del DNA

La metilació del DNA es dona a la posició 5 del carboni de les citosines, principalment en els dinucleòtids citosina-fosfat-guanidina (CpG) (Rousseaux *et al.*, 2005; Auclair and Weber, 2012). Es troba regulada pels enzims metiltransferases (DNMTs) on els enzims *DNA (cytosine-5)-methyltransferase 3A* (DNMT3A) i *DNA (cytosine-5)-methyltransferase 3B* (DNMT3B) són els encarregats de la metilació *de novo* i l'enzim *DNA (cytosine-5)-methyltransferase 1* (DNMT1) és l'encarregat del manteniment d'aquestes metilacions (Bestor, 2000). La metilació del DNA és un inhibidor de la transcripció, sobretot en les illes CpG que es troben a les regions promotores dels gens i afavoreixen o inhibeixen, dependent del seu estat hipo- o hiper- metilat, respectivament, la unió de factors de transcripció i altres proteïnes d'unió al DNA. Després de la fertilització, el genoma patern és activament desmetilat en la seva majoria (Veure apartat 1.3.2) tot i que queden regions metilades de gens tan d'*imprinting* parental com altres gens i retrotransposons (revisat a (Carrell *et al.*, 2016)). En l'espermatozoide, s'ha descrit que existeix una hipometilació general en els promotors de gens importants en el desenvolupament embrionari (Arpanahi *et al.*, 2009; Hammoud *et al.*, 2009).

1.4.3.2. Estructura de la cromatina espermàtica

Durant les darreres etapes de l'espermatogènesi, concretament durant l'espermiogènesi, les histones que empaqueten el DNA són seqüencialment reemplaçades, primerament per variants d'histones, seguit per proteïnes de transició i finalment per protamines (Oliva and Dixon, 1991; Oliva, 2006; Jodar and Oliva, 2014). Aquest procés, que requereix de variants d'histones i de PTMs de les histones (Veure annex 1), s'anomena transició nucleohistona – nucleoprotamina (NH – NP) i dona lloc a una dràstica reorganització on es passa d'una distribució nucleosomal basada en histones de la cèl·lula diploide a una distribució toroïdal basada en les protamines en l'espermatozoide madur haploide (Figura 1.10). Així doncs, aproximadament del 85-95 % del DNA de l'espermatozoide madur es troba unit a protamines (en humà ≈ 92 %), mentre que el 5-15 % del DNA restant roman unit a histones (en humà ≈ 8 %) (Gatewood *et al.*, 1987). Per tant, la cromatina espermàtica

presenta una estructura única formada per dos dominis amb nivells d'empaquetament diferents: el domini NH on el DNA es troba més accessible i el domini NP on el DNA es troba altament empaquetat en forma de toroides (Oliva, 2006; Castillo *et al.*, 2014a; Jodar and Oliva, 2014).

1.4.3.2.1. Domini nucleohistona (NH)

L'estructura dels nucleosomes a l'espermatozoide madur és similar a la de la cèl·lula somàtica. Consisteix en 147 bp de DNA enrotllat al voltant del nucli d'histona octamèrica que inclou dues molècules de *histone H2A type 2-A* (H2A), dues molècules de *histone H2B type 1* (H2B), dues molècules de h*istone H3.1* (H3) i dues molècules de *histone H4* (H4). La *histone H1* (H1) es troba enllaçant l'entrada i sortida del DNA de l'octàmer. A més, la H1 també juga un paper important en la modulació del plegament de la cromatina. Els nucleosomes adjacents estan interconnectats per DNA que pot arribar a tenir una longitud d'uns 80 bp (Ausió, 2015). El resultat final és un DNA 5 vegades més empaquetat, tot i que aquest empaquetament permet una disposició més oberta i dinàmica, que el conferit pel domini NP. Per tant, pot ser regulada i modulada a través de la incorporació de variants d'histones, PTMs d'histones i factors nuclears (Mezquita, 1985; Oliva and Dixon, 1991; Zhao *et al.*, 2004; Kimmins and Sassone-Corsi, 2005; Krejčí *et al.*, 2015; Luense *et al.*, 2016).

Durant l'espermatogènesi, en la transició NH – NP, les histones canòniques són reemplaçades per variants d'histones que desestabilitzen les interaccions proteïna DNA i permeten la transició a NP (Figura 1.10) (Kimmins and Sassone-Corsi, 2005; Oliva, 2006; Maze *et al.*, 2014). Encara que les variants d'histones només presenten canvis lleus en la seva estructura primària respecte les histones canòniques, aquestes diferències poden afectar àmpliament la seva estructura, estabilitat i funcionalitat (Govin *et al.*, 2004). Hi ha variants d'histones que s'expressen en altres tipus cel·lulars, però n'hi ha algunes que són específiques de testicle (Veure annex 1). L'eliminació específica en models animals (KO) de variants d'histones és de gran ajuda per determinar el seu paper en l'espermatogènesi i la fertilitat. A l'annex 1 es detallen totes les variants d'histones identificades en l'espermatozoide humà madur, la localització on s'expressen i les alteracions tan a nivell reproductiu com l'afectació que provoquen al desenvolupament embrionari de les variants en què s'ha realitzat un KO. Per exemple, l'absència de la variant *testis-specific H1 histone* (H1t2), provoca entre altres, una alteració en la fertilització (Martianov *et al.*, 2005; Tanaka *et al.*, 2005). També hi ha alguns KO de variants d'histona que provoquen letalitat embrionària, però no hi ha suficient evidència de que la causa sigui de la línia germinal (Veure annex 1).

Les histones i les seves variants incorporen PTMs. Les PTMs de les histones han estat àmpliament estudiades i s'ha demostrat que diferents combinacions de PTMs creen un codi complex, anomenat el codi de les histones, que contribueix a l'expressió gènica i a l'organització i dinàmica de la cromatina (Kimmins and Sassone-Corsi, 2005; Krejčí *et al.*, 2015). Està demostrat que l'acetilació de la H4 durant l'espermatogènesi és necessària per reduir l'afinitat dels nucleosomes pel DNA, relaxa la cromatina i permet el reemplaçament d'histones per protamines (Figura 1.10; veure annex 1) (Oliva *et al.*, 1987; Oliva, 2006). La transició NH – NP no té lloc totalment, sinó que una petita proporció del DNA de l'espermatozoide madur, resta empaquetat amb histones (Figura 1.10), en les quals s'han descrit múltiples PTMs diferents, incloent acetilació, metilació i fosforilació (Krejčí *et al.*, 2015). Aquestes histones amb les seves PTMs són aportades a l'embrió durant la fertilització i s'ha proposat que podrien tenir un paper rellevant en el control epigenètic del desenvolupament embrionari (Van Der Heijden *et al.*, 2008; Carrell and Hammoud, 2009). Molts estudis han suggerit que en l'espermatozoide madur, el romanent de DNA unit a histones no es troba distribuït aleatòriament, sinó que es troba associat a telòmers i centròmers i enriquit en gens importants pel desenvolupament embrionari i en la regulació de l'expressió gènica, incloent proteïnes de vies de senyalització i factors de transcripció involucrats en l'embriogènesi, gens amb *imprinting* parental i clústers de microRNAs (miRNAs) (Arpanahi *et al.*, 2009; Hammoud *et al.*, 2009; Vavouri and Lehner, 2011; Castillo *et al.*, 2014a, 2015; Siklenka *et al.*, 2015; Royo *et al.*, 2016). A nivell proteic, es va determinar que la fracció NH conté, a part de variants d'histones, factors de transcripció, proteïnes associades a la cromatina amb rols reguladors, també suggerint el potencial epigenètic aportat per l'espermatozoide en el moment de la fertilització (Castillo *et al.*, 2014a). Per tant, no només el contingut d'histones, variants i PTMs aporta informació a l'embrió en el moment de la fertilització, sinó que la distribució no aleatòria NH - NP també estaria aportant informació epigenètica al zigot i podria regular l'expressió gènica durant el desenvolupament preimplantacional (Veure annex 1).

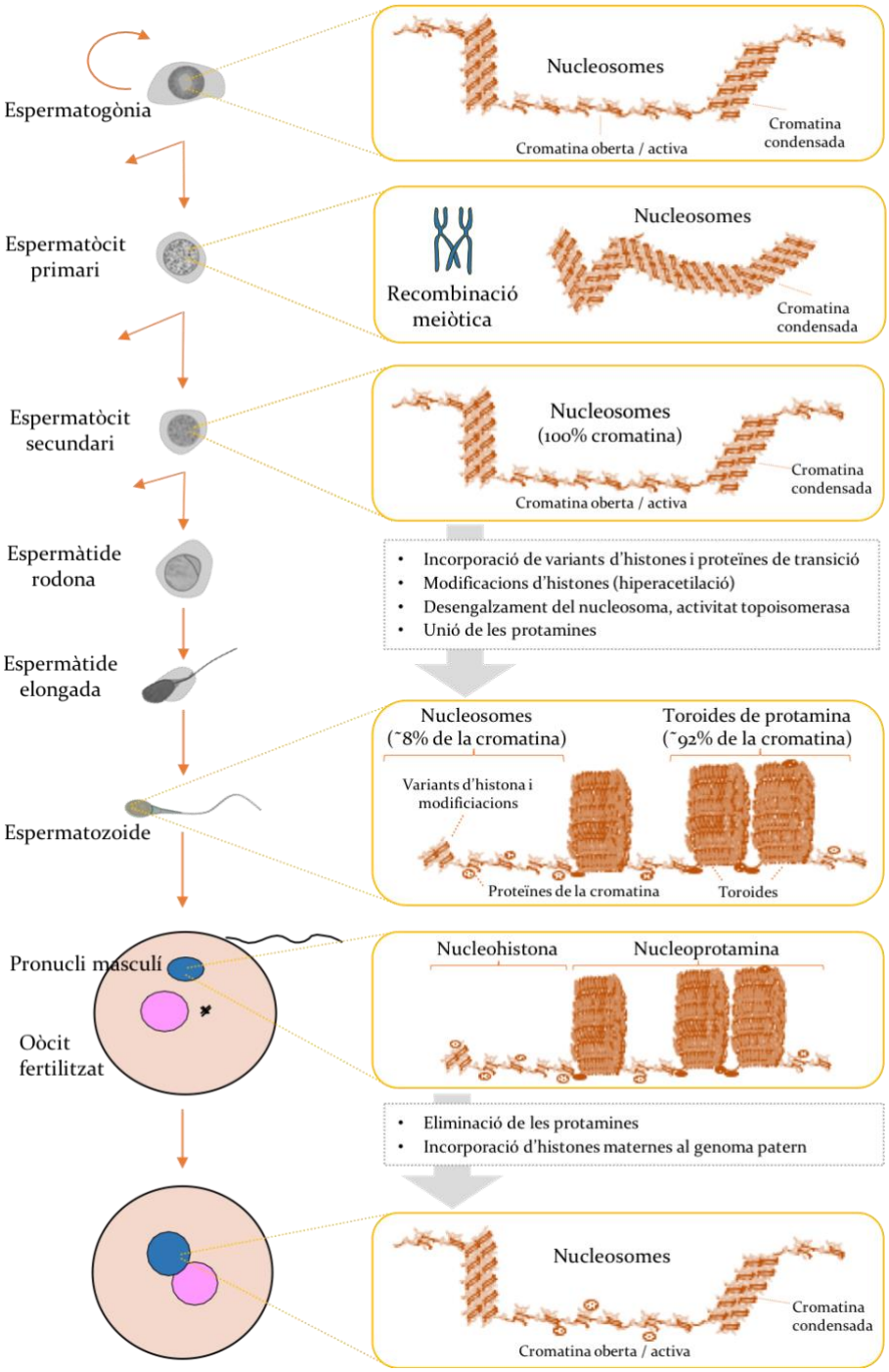


Figura 1.10 Canvis en la cromatina durant l'espermatogènesi i la fertilització. Un cop l'espermatogònia es diferencia a espermatòcit primari, té lloc la recombinació meiòtica i seguidament dona lloc a l'espermatòcit secundari. L'espermatòcit secundari es dividirà i donarà lloc a l'espermàtida rodona que serà haploide. Fins aquesta etapa, la cromatina es troba organitzada com una cèl·lula somàtica on el DNA es troba empaquetat amb nucleosomes. És durant el procés d'espermiogènesi quan les histones són substituïdes primerament per variants d'histones, proteïnes de transició i finalment per protamines, amb l'ajuda de les topoisomerases. En humans, el DNA de l'espermatozoide madur es troba empaquetat en un aproximadament 92 % amb protamines i un 8 % amb nucleosomes. Així doncs, l'espermatozoide madur presenta dos dominis clarament diferenciats, el domini nucleohistona i el domini nucleoprotamina que seran transferits a l'oòcit en el moment de la fertilització. Després de la fertilització, les protamines són substituïdes per histones maternes i retorna a una distribució totalment nucleosomal. Adaptat de l'Annex 1 (Barrachina *et al.*, 2018).

1.4.3.2.2. Domini nucleoprotamina (NP)

Les protamines són les proteïnes més abundants del nucli de l'espermatozoide i com s'ha mencionat anteriorment, empaqueten el 85-95 % del DNA patern (Figura 1.10) (Oliva and Dixon, 1991; Oliva, 2006; Balhorn, 2007; Castillo *et al.*, 2014a; Jodar and Oliva, 2014; Barrachina *et al.*, 2018). Les protamines són unes proteïnes petites, d'uns 50-102 aminoàcids, molt bàsiques i riques en residus arginina que es troben carregats positivament i que permeten la forta unió amb el DNA carregat negativament. També són riques en residus cisteïna (Figura 1.11), els quals permeten la formació de ponts disulfur i ponts de Zn^{2+} intra- i inter- protamina (Balhorn *et al.*, 1992; Bench *et al.*, 2000; Oliva, 2006; Björndahl and Kvist, 2010). Es creu que la formació de ponts de salts, com els ponts de Zn^{2+}, tenen la funció d'estabilitzar la cromatina (Kvist and Björndahl, 1985; Björndahl and Kvist, 2010). El resultat final és la formació del complex NP en forma de toroide que empaqueta fins a 50 kb de DNA (Figura 1.10; veure annex 1) (Oliva, 2006).

En mamífers s'han descrit dos tipus de protamina, la protamina 1 (P1) i la família de la protamina 2 (P2), tot i que P2 només és expressada a nivell proteic per algunes espècies com són ratolí i humà (Balhorn *et al.*, 1987; Oliva, 1995, 2006; Queralt *et al.*, 1995). En humans, els gens codificants per la protamina 1 (*PRM1)* i per la protamina 2 (*PRM2)* es troben situats al cromosoma 16p13.3 (Nelson and Krawetz, 1995). Mentre que la P1 és sintetitzada directament a la proteïna madura de 51 aminoàcids, la família de P2 és generada a partir del precursor de 102 aminoàcids (pre-P2; Figura 1.11; (de Mateo *et al.*, 2011b)). Per proteòlisi pre-P2 dona lloc als 3 components madurs de la família de P2, HP2, HP3 i HP4, que només difereixen entre ells de 1-4 aminoàcids al seu extrem N-terminal (Figura 1.11). La forma madura HP2 és la més abundant, seguit de HP3 i HP4 (Oliva, 2006; Jodar and Oliva, 2014). Alguns estudis, també, han identificat formes intermèdies processades de pre-P2, anomenades b*asic nuclear protein*s HPI1, HPI2, HPS1 i HPS2 (Figura 1.11) que podrien donar lloc a les formes madures de P2 (de Mateo *et al.*, 2011b). Un cop té lloc la fertilització de l'oòcit, les protamines són ràpidament substituïdes per histones maternes (Jones *et al.*, 2012).

S'han proposat diverses funcions per a les protamines. Entre altres: (i) empaquetar fortament el DNA patern per a què el nucli de l'espermatozoide sigui més compacte i dinàmic, necessari per a la motilitat espermàtica; (ii) protegir el DNA espermàtic de mutàgens i nucleases presents en els tractes sexuals masculí i femení; (iii) competir i eliminar proteïnes nuclears associades a la cromatina durant l'espermatogènesi per bloquejar la transcripció i la traducció; i (iv) participar en *l'imprinting* del genoma patern durant l'espermatogènesi i conferir marques epigenètiques en algunes regions del genoma que afecten la seva reactivació després de la fertilització (Veure annex 1) (Oliva, 2006).

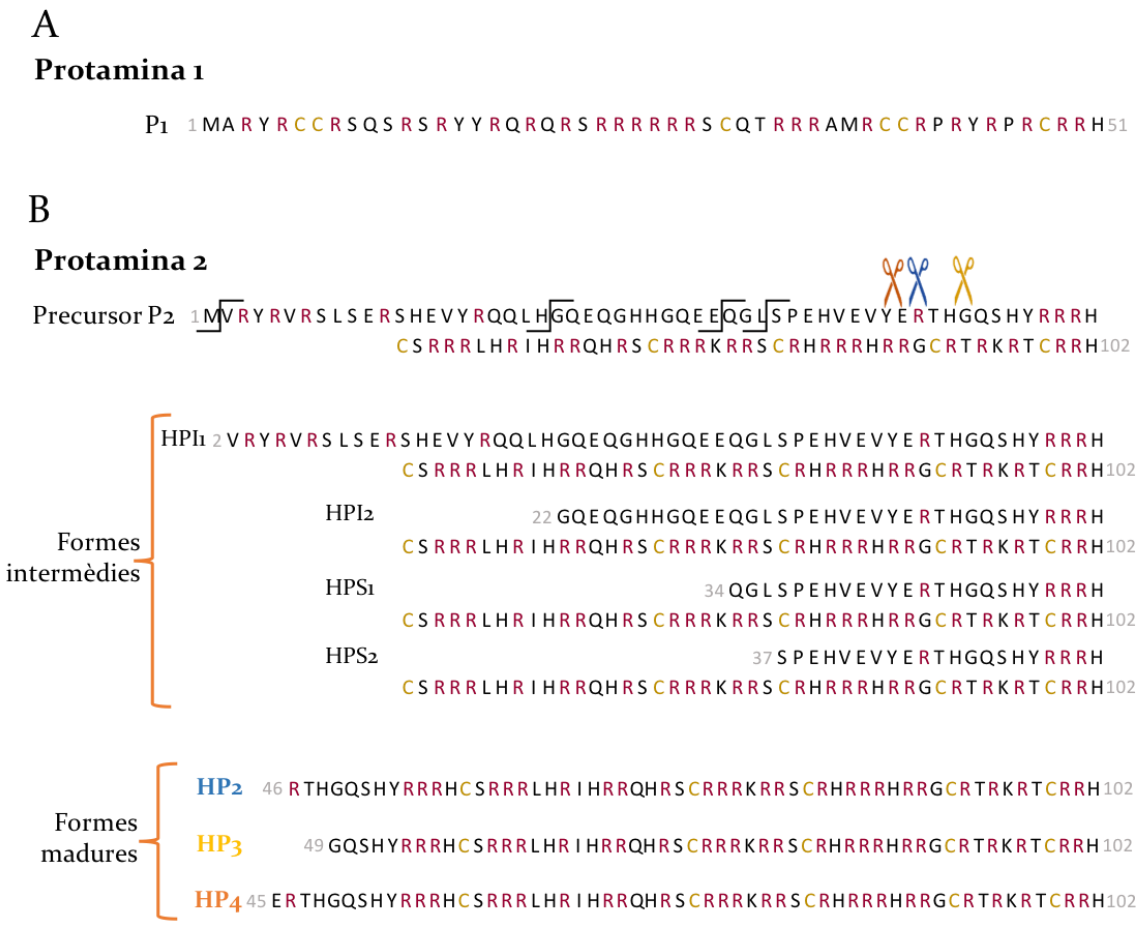


Figura 1.11 Seqüència d'aminoàcids de les protamines. A) Seqüència de la protamina 1 (P1). B) Seqüència del precursor de la protamina 2 (pre-P2), les formes intermèdies (HPI1, HPI2, HPS1 i HPS2) i les formes madures que s'obtenen per la proteòlisi de pre-P2 (HP2, HP3 i HP4). En vermell estan indicats els residus arginina i en groc ataronjat els residus cisteïna. Les línies negres marquen el lloc de proteòlisi de les formes intermèdies. Les tisores marquen el lloc de proteòlisi de les formes madures HP2 (blau), HP3 (groc); HP4 (taronja). Basat en (Oliva, 2006; de Mateo *et al.*, 2011b; Jodar and Oliva, 2014).

Existeixen vàries tècniques clàssiques per a calcular la quantitat de protamines i possibles alteracions. Aquestes es poden calcular a través de la quantificació dels nivells d'RNA missatger (mRNA) de les protamines, a través de tècniques de tinció indirecta de la cromatina o bé a través de la quantificació directa (Nasr-Esfahani *et al.*, 2004b; Aoki *et al.*, 2006b; Oliva, 2006; Jodar and Oliva, 2014; Ribas-Maynou *et al.*, 2015; Ni *et al.*, 2016). Tot i així, la tècnica més àmpliament utilitzada és la quantificació directa de protamines a nivell proteic, que es basa en l'extracció específica àcida de les protamines i la seva visualització a través de l'ús de tècniques convencionals d'electroforesi en gels de poliacrilamida àcid – urea (àcid – urea PAGE), del qual es calcula el ràtio relatiu entre l'abundància de P1 i els membres de la família de P2 (P1/P2) i s'utilitza com a mesura de la maduresa de la cromatina i se n'estableix la normalitat / alteració (Oliva, 2006).

El ràtio normal entre la quantitat de proteïna de P1 i de la família de P2 es va establir al voltant de 1 (0,8 – 1,2) (Balhorn *et al.*, 1988; Corzett *et al.*, 2002). La desregulació de les protamines ha estat lligada a defectes en el desenvolupament embrionari tan a nivell d'RNA com de proteïna. Concretament, a nivell d'RNA, s'han relacionat alteracions en els transcrits de *PRM1* i *PRM2* amb una disminució en la capacitat de fertilització i amb avortaments espontanis (Depa-Martynów *et al.*, 2007; Steger *et al.*, 2008; Depa-Martynow *et al.*, 2012; Rogenhofer *et al.*, 2013, 2017). A nivell de proteïna, s'ha reportat en múltiples estudis que alteracions en el ràtio P1/P2 i el ràtio pre-P2/P2 estarien associades amb baixes taxes de fertilització i/o implantació, una baixa qualitat embrionària i una baixa taxa d'embaràs utilitzant tan FIV convencional com FIV-ICSI (Khara, 1997; Nasr-Esfahani *et al.*, 2004b, 2004a; Aoki *et al.*, 2005a, 2006b; de Mateo *et al.*, 2009; Simon *et al.*, 2011). Per tant, aquests resultats posen de manifest la possible implicació de les protamines amb la fertilització i el desenvolupament embrionari primerenc.

1.4.4. RNAs

Tot i que en l'espermatozoide madur la transcripció i la traducció es troben aturades a causa de l'alta compactació de la cromatina i l'eliminació de la major part del citoplasma que conté la maquinària traduccional, el gàmeta masculí conté una complexa població d'RNAs retinguts selectivament a la cèl·lula espermàtica (Jodar *et al.*, 2013). Utilitzant la tècnica heteròloga on espermatozoides humans podien fecundar oòcits de ratolí que no presentaven zona pel·lúcida, es va determinar que l'espermatozoide aportava un *cargo* d'RNAs en el moment de la fertilització que podrien tenir un paper important en el desenvolupament embrionari primerenc (Ostermeier *et al.*, 2002, 2004; Jodar *et al.*, 2013). Aquest possible rol es reflecteix en què, en ratolí, quan s'injecten en oòcits espermatozoides tractats prèviament que contenen menys del 90 % d'RNAs que presenten normalment, disminueix el percentatge d'embrions que arriben a l'estadi de blastocist i també disminueix el nombre de naixements; mentre que quan s'injecta la totalitat d'RNAs d'espermatozoides no tractats en aquests embrions deficients, el fenotip es rescata (Guo *et al.*, 2017). També, s'ha descrit que els RNAs de l'espermatozoide semblen ser claus en la consolidació / confrontació dels genomes. Quan l'espermatozoide i l'oòcit entren en contacte és molt important assegurar-se de la seva compatibilitat i de la correcta contribució de cada part per a què un cop es combinin tingui lloc un correcte desenvolupament embrionari (Patil and Totey, 2003; Bourc'his and Voinnet, 2010; Goring and Indriolo, 2010; Krawetz *et al.*, 2011; Jodar *et al.*, 2013; Miller and Iles, 2013; Miller, 2015; Gòdia *et al.*, 2018).

1.4.4.1. RNAs codificants

Arran de que l'espermatozoide és una cèl·lula que no transcriu ni tradueix, es pensava que els RNAs codificants presents a l'espermatozoide madur eren romanents de l'espermatogènesi sense cap funció, però s'ha demostrat que hi ha una població estable d'RNAs codificants intactes a la cèl·lula espermàtica, la qual cosa indica la seva retenció específica durant l'espermatogènesi (Ostermeier *et al.*, 2002, 2004, 2005).

Un dels mecanismes proposats pels RNAs codificants de l'espermatozoide que es conserven intactes en el moment de la fertilització és la seva traducció mitjançant la maquinària traduccional activa de l'oòcit fecundat (Jodar *et al.*, 2013; Miller, 2015; Ntostis *et al.*, 2017; Castillo *et al.*, 2018; Jodar, 2019). Per exemple, l'RNA *integrator complex subunit 1* (*INTS1*), que participa en la transcripció i el processament de diferents *small-nuclear RNAs* (snRNAs), es troba intacte i retingut a l'espermatozoide madur i per *microarray* s'ha determinat que incrementa després de la fertilització abans de l'activació del genoma embrionari. L'absència específica de la proteïna INTS1 en ratolins provoca letalitat embrionària en l'estadi de blastocist (Hata and Nakayama, 2007; Vassena *et al.*, 2011; Jodar *et al.*, 2013). Podria ser que aquest RNA codificant aportat per l'espermatozoide fos traduït per la maquinària oocitària i participés en el desenvolupament embrionari primerenc. De la mateixa manera, s'ha demostrat en ratolí, que la injecció del transcrit de *Plcz1*, essencial per l'inici del desenvolupament embrionari (Veure apartat 1.3) és traduït i provoca l'activació de l'oòcit *in vitro* (Yoda *et al.*, 2004). Per tant, aquests resultats posen en evidència la possible contribució dels RNAs codificants de l'espermatozoide en els primers estadis del desenvolupament embrionari primerenc.

1.4.4.2. RNAs no codificants

Com s'ha vist en els RNAs codificants (Veure apartat 1.4.4.1), també s'ha postulat que alguns RNAs no codificants (ncRNAs) poden tenir certa funció en el desenvolupament embrionari primerenc.

Aproximadament el 7 % dels *small* ncRNAs de l'espermatozoide humà són miRNAs i són els RNAs no codificants que més s'han estudiat (Gòdia *et al.*, 2018). Regulen l'expressió gènica inhibint o activant la traducció dels seus mRNAs diana o bé unint-se a la regió 3' *untranslated region* (UTR) dels mRNAs diana per provocar-ne la seva degradació. El precursor del precursor del miRNA (pri-miRNA) és sintetitzat per la RNA polimerasa II i processat per la *ribonuclease 3* (DROSHA) i el *DGCR8 microprocessor complex subunit* (DGCR8) fins esdevenir precursor del miRNA (pre-miRNA). Seguidament el pre-miRNA és exportat al citoplasma on és processat altre cop pel complex *endoribonuclease Dicer* (DICER1) i *human immunodeficiency virus transactivating response RNA-binding protein* (TRBP), donant lloc a la forma madura

de 22 – 24nt de longitud (Jodar *et al.*, 2013). En un estudi en model murí, s'ha observat que el KO *Drosha*$^{-/-}$ o *Dicer*$^{-/-}$ condicionals de testicle, implicats en la síntesi dels miRNAs i de *small interfering* RNAs (siRNAs), presenten un perfil alterat de miRNA i siRNAs que provoca una reducció en el potencial de desenvolupament embrionari (Yuan *et al.*, 2016). No obstant, el fenotip WT es rescatava quan s'injectava RNAs d'espermatozoides de ratolins WT als embrions provinents dels KO, posant de manifest la rellevància dels miRNAs i siRNAs en el desenvolupament embrionari (Yuan *et al.*, 2016). Un altre exemple és el miR-34c, que es va proposar que era necessari per la primera divisió embrionària i per la correcta embriogènesi, però existeix certa controvèrsia (Liu *et al.*, 2012; Yuan *et al.*, 2015; Shi *et al.*, 2020). També s'ha proposat que certs pri-miRNAs aportats per l'espermatozoide en el moment de la fertilització, com el pri-miRNA-181c, el qual les seves dianes es troben en gens del desenvolupament, podrien ser processats a la seva forma madura per la maquinària materna i realitzar la seva funció abans de l'activació del genoma embrionari (Vassena *et al.*, 2011; Jodar *et al.*, 2013; Sendler *et al.*, 2013).

Un altre tipus d'RNA molt abundant en l'espermatozoide humà són els *P element-induced wimpy testes-interacting (piwi-interacting) RNAs* (piRNAs) que es calcula que conformen un 17 % del total de *small* ncRNAs (Krawetz *et al.*, 2011; Pantano *et al.*, 2015). Els piRNAs s'organitzen al llarg del genoma en forma de clúster d'entre 1 – 100 Kb. La seva biogènesi encara no està del tot coneguda tot i que és dependent de la proteïna Piwi (PIWI) i després de ser processats, la forma madura conté entre 23 – 32 nt. Durant l'espermatogènesi els elements transposables són inhibits per piRNAs i s'ha demostrat que el dèficit de certs piRNAs durant l'espermatogènesi provoquen defectes en l'espermatozoide i redueixen la seva capacitat de fertilització (Wu *et al.*, 2020). Una de les funcions proposades pels piRNAs aportats per l'espermatozoide madur en el moment de la fertilització és que podrien estar associats a protegir el genoma durant els primers estadis del desenvolupament embrionari (Jodar *et al.*, 2013; Pantano *et al.*, 2015).

Similarment, els elements transposables que alineen a regions repetitives del genoma podrien participar en el desenvolupament embrionari. Els més representats són membres de les famílies *long terminal repeat* (LTR), *short interspersed nuclear element* (SINE), repeticions ALU i *long interspersed nuclear element* (LINE) (Krawetz *et al.*, 2011). Tot i que el paper que juguen els elements transposables al desenvolupament embrionari encara és controvertit (Jodar *et al.*, 2013), s'ha vist per exemple que la inactivació del *long interspersed nuclear element* 1 (*LINE1*) a l'embrió causa un arrest a l'estadi de 2 – 4 cèl·lules, reflectint la seva importància pel correcte desenvolupament (Beraldi *et al.*, 2006). Més recentment, s'ha associat alteracions en l'activitat de LINE1 a avortaments de repetició (Lou *et al.*, 2020).

1.4.5. Proteïnes

L'espermatozoide aporta en el moment de la fertilització un conjunt de proteïnes, algunes d'elles implicades en la regulació de l'expressió gènica, que podrien tenir un paper clau en la fertilització i el desenvolupament embrionari primerenc (Castillo *et al.*, 2018). Per exemple, l'absència de la proteïna *sperm equatorial segment protein 1* (Spesp1) en ratolins, provoca una reducció de la capacitat fertilitzant de l'espermatozoide, ja que afecta la localització i quantitat de proteïnes implicades en la fusió com són Izumo1 i la família de proteïnes *disintegrin and metalloproteinase domain-containing protein* (Adam) (Fujihara *et al.*, 2010; Castillo *et al.*, 2018). La generació de ratolins KO per la proteïna espermàtica *proprotein convertase subtilisin/kexin type-4* (Pcsk4), són incapaços de processar la proteïna *acrosin binding protein* (Acrbp). Aquest fet implica una capacitació accelerada, una reacció acrosòmica precipitada i la incapacitat de l'espermatozoide a unir-se a la zona pel·lúcida, impedint que tingui lloc la fertilització (Gyamera-Acheampong *et al.*, 2006; Tardif *et al.*, 2012; Castillo *et al.*, 2018). Similarment, la proteïna de l'espermatozoide CRISP1, que pertany a la família de proteïnes CatSper, és incorporada durant la maduració posttesticular a l'epidídim i es troba relacionada amb la penetració i fusió de l'espermatozoide a l'oòcit, específicament, marca l'orientació d'entrada al cúmulus i interactua amb la zona pel·lúcida. S'ha descrit que ratolins *Crisp1*$^{-/-}$, són incapaços de penetrar i fertilitzar l'oòcit (Busso *et al.*, 2007; Cohen *et al.*, 2011; Maldera *et al.*, 2014; da Ros *et al.*, 2015; Castillo *et al.*, 2018). De manera més extrema, s'ha especulat que l'absència específica d'alguna proteïna espermàtica també podria provocar letalitat embrionària (Castillo *et al.*, 2018).

1.5. Infertilitat masculina

La infertilitat es defineix com la incapacitat d'una parella en aconseguir embaràs després d'un any mantenint relacions sexuals regulars sense l'ús de mètodes anticonceptius (Evers, 2002; World Health Organization, 2010). Tot i que generalment es considera sinònim parlar d'infertilitat i esterilitat, alguns autors consideren que l'esterilitat és la incapacitat de concebre de manera irreversible, mentre que la infertilitat és la reducció de la capacitat de concebre. Arran dels avenços en les tècniques de reproducció assistida (TRAs), els límits irreversibles s'han anat superant i s'ha definit el terme subfertilitat per a designar totes les persones que no aconsegueixen tenir descendència de manera espontània després d'un any mantenint relacions sexuals sense protecció durant la finestra fèrtil del cicle menstrual (Evers, 2002). Per tant, se sol utilitzar infertilitat i subfertilitat de manera sinònima. La infertilitat és un problema que ateny al 15 – 20 % de les parelles en edat fèrtil. Aproximadament d'aquestes, un 40 % dels casos són deguts a un factor exclusivament femení, un 40 % deguts a un factor exclusivament masculí i un 20 % on ambdós membres de la parella presenten alteracions reproductives (Ballescà and Oliva, 2012). Així doncs, el factor masculí pot arribar a contribuir en el 50 % de causes de infertilitat (Vander Borght and Wyns, 2018).

Al llarg del temps, les causes de la infertilitat s'han centrat majoritàriament en l'estudi de la dona, a les quals se li realitzen múltiples proves exhaustives per determinar la possible etiologia de la subfertilitat, mentre que a l'home, l'estudi que se li realitza sol ser molt més superficial i limitat majoritàriament a l'estudi del semen (Cahill and Wardle, 2002; McLachlan *et al.*, 2005).

1.5.1. Causes conegudes de la infertilitat masculina

Per a què un home sigui fèrtil ha de desenvolupar-se de manera normal l'espermatogènesi, la maduració i emmagatzematge dels espermatozoides a l'epidídim, l'adquisició dels diferents fluids de les glàndules sexuals accessòries durant l'ejaculació i, finalment, les relacions sexuals han de desenvolupar-se satisfactòriament. Per aquest motiu, la infertilitat masculina es pot classificar com a congènita, adquirida o idiopàtica (Krausz, 2011). A la Taula 1.1 es detallen algunes de les causes conegudes de la infertilitat masculina.

Existeixen factors exògens que provoquen infertilitat com ara l'ús de fàrmacs citotòxics, radiació, calor o estil de vida; alteracions físiques degut a defectes congènits, obstruccions, traumatismes o cirurgies; varicocele; infeccions testiculars, epididimàries o de les glàndules sexuals accessòries; alteracions endocrines com l'hipogonadisme-hipogonadotròpic; causes immunològiques; o causes

psicològiques. També, s'han descrit múltiples alteracions genètiques que provoquen infertilitat com són la Síndrome de Klinefelter (47, XXY), les microdelecions del cromosoma Y, la Síndrome de Kallmann o mutacions puntuals en diferents gens (Taula 1.1) (de Kretser, 1997; Niederberger, 2009; Krausz, 2011; Krausz and Riera-Escamilla, 2018). Tot i els avenços en el camp de la infertilitat masculina, aproximadament el 50 % dels casos continuen essent sense causa coneguda (Krausz, 2011).

Taula 1.1. Classificació de les possibles causes de la infertilitat masculina. Adaptat de (Krausz, 2011).

Tipus de factor	Causa
Factors congènits	Alteracions genètiques (Síndrome de Klinefelter, microdelecions del cromosoma Y, Síndrome de Kallmann, mutacions en gens involucrats en l'eix Hipotàlem-Hipòfisi-gonadal, ...)
	Anòrquia / Criptorquídia
	Absència congènita del Vas deferens
Factors adquirits	Malalties sistèmiques
	Factors endocrins o hipogonadisme-hipogonadotròpic adquirit
	Trauma / torsió testicular
	Causes post-inflamatòries (orquitis, epididimitis)
	Obstrucció / subobstrucció del tracte urogenital
	Infeccions urogenitals recurrents
	Disfunció erèctil o ejaculatòria
	Cirurgia
	Varicocele
	Factors exògens (medicació, fàrmacs citotòxics, radiació, calor)
	Estil de vida
Factors idiopàtics	Causa desconeguda

1.5.2. Anàlisi dels paràmetres seminals

Tot i que existeixen diverses tècniques per avaluar l'estat fèrtil d'un individu com ara l'exploració física i la determinació dels nivells hormonals (Revisat a (Niederberger, 2009; Pan *et al.*, 2018)), una de les tècniques més àmpliament utilitzades és la l'avaluació dels paràmetres seminals a través de la realització d'un seminograma. En el seminograma es podrà valorar a nivell macroscòpic el pH i el volum de la mostra i a nivell microscòpic la quantitat, la motilitat, la vitalitat i la morfologia dels espermatozoides. Existeix el manual realitzat per la *World Health Organization* (WHO) on es detalla com realitzar i interpretar un seminograma (World Health Organization, 2010) i on s'estableixen els valors de referència d'una mostra de semen (Taula 1.2). Per poder realitzar un diagnòstic del tipus d'infertilitat, almenys s'han d'obtenir 2 mostres obtingudes en uns 3 mesos de diferència, que corresponen a dos cicles d'espermatogènesi diferents (Niederberger, 2009; World Health Organization, 2010).

Taula 1.2. Valors de referència mínims dels paràmetres seminals avaluats en un seminograma establerts per la *World Health Organization* (WHO). Adaptat de (World Health Organization, 2010).

Paràmetre	**Límit inferior de referència**
Volum del semen (ml)	1.5 (1.4 – 1.7)
pH	≥ 7.2
Nombre total d'espermatozoides (10^6 per ejaculat)	39 (33 – 46)
Concentració espermàtica (10^6 per ml)	15 (12 – 16)
Motilitat total (progressiva + no progressiva, %)	40 (38 – 42)
Motilitat progressiva (%)	32 (31 – 34)
Vitalitat (espermatozoides vius, %)	58 (55 – 63)
Morfologia espermàtica (formes normals, %)	4 (3 – 4)

La mostra s'obté per masturbació en un contenidor estèril després d'un període d'abstinència de 3 – 7 dies (World Health Organization, 2010). Un cop la mostra ha estat obtinguda es deixa liquar almenys durant 30 minuts abans de ser analitzada. Mitjançant un microscopi amb contrast de fases, es determina la possible presència de cèl·lules contaminants, com són els leucòcits, les cèl·lules germinals immadures o cèl·lules epitelials de descamació (World Health Organization, 2010). Per analitzar els paràmetres seminals de la mostra de semen es pot realitzar o bé mitjançant sistemes automatitzats de comptatge anomenats *Computer-Assisted Sperm Analysis* (CASA) o bé mitjançant sistemes manuals amb l'ús de càmeres de comptatge com l'hemocitòmetre de Neubauer o la cambra de Makler (World Health Organization, 2010).

Principalment en el seminograma s'analitzen els següents paràmetres:

- **Concentració:** Es determina el nombre exacte de milions d'espermatozoides per mil·lilitre de semen (10^6 / ml) i també es calculen els milions d'espermatozoide totals per ejaculat (World Health Organization, 2010). Els valors de referència per considerar una mostra normal són ≥ 15 milions / ml de semen i ≥ 39 milions totals per ejaculat (Taula 1.2).

- **Motilitat:** La motilitat espermàtica es pot dividir en 4 subtipus i es considera normal si existeix ≥ 32 % de motilitat progressiva o ≥ 40 % de motilitat total (Taula 1.2):
    - **Espermatozoides amb mobilitat progressiva (Tipus A+B)**: Són espermatozoides que es mouen activament de manera lineal o realitzant un cercle gran, independentment de la velocitat. Anteriorment, es classificaven els espermatozoides progressius en mobilitat ràpida (A) i lenta (B), on els espermatozoides de tipus A

es movien a > 25 metres / segon (m / s) a 37 °C. Tot i així, degut a la dificultat de diferenciar la velocitat ràpida i lenta, es va deixar d'utilitzar (Cooper and Yeung, 2006).

- **Espermatozoides amb motilitat no progressiva (Tipus C):** Són la resta de patrons de mobilitat que s'observen en absència de progressió, com ara quan s'observa només el batec flagel·lar, quan els espermatozoides neden en petits cercles o quan el flagel només mou lleument el cap.
- **Espermatozoides immòbils (Tipus D):** Quan l'espermatozoide no presenta cap mena de moviment.

- **Morfologia:** Per valorar la morfologia espermàtica, els espermatozoides es fixen i es tenyeixen amb la tinció de *Diff-Quick, Papanicolau* o *Shorr* (World Health Organization, 2010). Per considerar un espermatozoide normal el cap ha de ser ovalat, amb una regió acrosòmica ben definida que en comprengui del 40 – 70 %. La regió acrosòmica no ha de contenir vacuoles grans i no hauria de contenir més de dues vacuoles petites, les quals ocupin més del 20 % del cap de l'espermatozoide. La regió post-acrosòmica no ha de contenir vacuoles. La peça intermèdia ha de ser regular i esvelta i ha de tenir mateixa longitud que el cap. També, han d'estar alineades la peça intermèdia i el cap de l'espermatozoide. La gota citoplasmàtica es considera anormal quan excedeix un terç del cap de l'espermatozoide (Mortimer and Menkveld, 2001). La peça principal ha de ser uniforme, més estreta que la peça intermèdia i ha de ser unes 10 vegades l'allargada del cap (World Health Organization, 2010). Una mostra de semen es considera que presenta morfologia normal si almenys el 4 % dels espermatozoides són normals (Taula 1.2). Els defectes dels espermatozoides es poden classificar en defectes de cap, defectes de coll i peça intermèdia o defectes de cua o la combinació de més d'una alteració. Se solen utilitzar els criteris estrictes descrits pel Dr. Kruger per determinar la morfologia (Kruger *et al.*, 1987).

- **Vitalitat:** S'estima en testar la integritat de la membrana dels espermatozoides, és a dir, els espermatozoides vius presenten la membrana íntegra mentre que els espermatozoides morts no. Per determinar la vitalitat es pot utilitzar un colorant com el test d'eosina o bé una solució hipotònica (World Health Organization, 2010). La vitalitat s'ha de realitzar just després de la liqüefacció o com a molt tard una hora després de l'ejaculació (World Health Organization, 2010). Una mostra normal ha de tenir una vitalitat d'almenys el 58 % dels espermatozoides (Taula 1.2).

Un cop el seminograma ha estat realitzat, es poden classificar els pacients en funció dels paràmetres seminals (World Health Organization, 2010).

- **Normozoospèrmia:** Tots els paràmetres seminals de la mostra es troben dins els rangs normals (Taula 1.2). Tot i així, que aparentment una mostra de semen presenti tots els paràmetres normals no vol dir que l'individu sigui fèrtil. De fet, la infertilitat d'origen desconegut en individus normozoospèrmics ha estat poc estudiada, tot i que suposa aproximadament del 6 % al 37 % de pacients infèrtils (Hamada *et al.*, 2012) i el seu estudi és de gran interès, ja que no és fàcil aconsellar quines TRAs seran més satisfactòries ni predir-ne les taxes d'èxit (Veure apartat 1.5.4).

- **Astenozoospèrmia:** Els espermatozoides presenten una alteració en la motilitat espermàtica, és a dir, presenten o bé < 32 % motilitat progressiva o bé < 40 % de motilitat total (Taula 1.2).

- **Teratozoospèrmia:** Els espermatozoides presenten una alteració en la morfologia espermàtica, és a dir, < 4 % dels espermatozoides presenten morfologia normal (Taula 1.2).

- **Oligozoospèrmia:** La concentració espermàtica i/o el nombre total d'espermatozoides es troba per sota de 15 milions / ml o 39 milions totals, respectivament (Taula 1.2).

- **Azoospèrmia:** Absència total d'espermatozoides després de centrifugar la mostra de semen. Pot ser deguda a una azoospèrmia obstructiva on hi ha alguna obstrucció a les vies seminals o bé a una azoospèrmia secretora on hi sol haver defectes en l'espermatogènesi (Taula 1.2).

- **Criptozoospèrmia**: Absència d'espermatozoides quan s'observa al microscopi però si se centrifuga la mostra, se n'observen al sediment.

Hi ha individus que presenten una combinació de diferents alteracions en els paràmetres seminals. Els individus que presenten baixa motilitat i alteració en la morfologia reben el nom d'astenoteratozoospèrmics, mentre que els individus que presenten baixa concentració i baixa motilitat reben el nom d'oligoastenozoospèrmics. Els individus que presenten baixa concentració i alteració en la morfologia reben el nom d'oligoteratozoospèrmics. Els individus que presenten la combinació de baixa concentració, baixa motilitat i alteració en la morfologia reben el nom d'oligoastenoteratozoospèrmics (World Health Organization, 2010).

1.5.3. Tècniques de reproducció assistida aplicades a la infertilitat masculina

Aproximadament el 2,1 % dels nens nascuts a Europa són gràcies a l'ús de les TRAs (de Geyter *et al.*, 2018). Les TRA són una opció per les parelles que no són capaces d'aconseguir embaràs de forma natural. Tot i així, aquest procés pot comportar un elevat cost econòmic i psicològic per a les parelles. La taxa d'èxit de les TRA se situa prop del 55 % després de 3 cicles i arriba al 85 % després de 8 cicles, per tant, són necessaris múltiples cicles abans d'aconseguir un naixement (Connolly *et al.*, 2010; Luke *et al.*, 2012). Un cop estudiats ambdós membres de la parella i dependent de les causes i el tipus d'infertilitat s'aplicarà una TRA específica. A grans trets, les TRA més comunes són: la inseminació intrauterina artificial (IIA), la FIV convencional i la FIV-ICSI.

- **Inseminació intrauterina artificial (IIA)**: Es tracta de la TRA menys invasiva i consisteix en dipositar els espermatozoides (prèviament seleccionats i capacitats al laboratori) a l'interior de l'úter de la dona (Figura 1.12A). Així doncs, aquesta tècnica mimetitza la concepció natural, però, permet que l'espermatozoide eviti l'ambient hostil del tracte sexual femení i el recorregut per fertilitzar l'oòcit és més curt. Tot i així, per a realitzar aquesta tècnica, són necessaris el major nombre possible d'espermatozoides mòbils, ja que només es facilita el trànsit inicial. S'aconsella en parelles amb subfertilitat d'origen desconegut, en parelles on existeix un factor femení lleu com malformacions en el coll uterí o problemes en l'ovulació o en parelles amb un factor masculí lleu.

- **Fecundació *in vitro* convencional (FIV convencional):** Aquesta tècnica es considera invasiva i es realitza la fecundació externament. Primerament, s'estimula hormonalment a la dona per a què maduri més oòcits que en un cicle natural. Seguidament, per punció s'extreuen aquests oòcits. En paral·lel, es prepara i capacita la mostra de semen, però en aquest cas és necessari utilitzar un nombre concret d'espermatozoides (uns 10^5), ja que han de ser suficients per a què trobin l'oòcit i el fertilitzin, però un excés podria provocar polispèrmia i, per tant, l'embrió no seria viable. Així doncs, en una placa amb el medi de cultiu adient es co-incuben els oòcits amb els espermatozoides per a què tingui lloc la fertilització (Figura 1.12B). Els embrions obtinguts són cultivats *in vitro* de 2 – 5 dies i per criteris de qualitat morfològica se'n seleccionaran de 1 a 3 per ser dipositats a l'úter de la dona, ja que la legislació Espanyola prohibeix la transferència de més de 3 embrions. Tot i que els criteris morfològics ajuden a predir la taxa d'èxit, no sempre hi ha una implantació satisfactòria. Aquesta tècnica s'aconsella en parelles amb

infertilitat d'origen desconegut on la IIA ha fallat i en parelles on hi ha un factor femení com ara endometriosi o obstrucció de les trompes de Fal·lopi. També es pot utilitzar quan hi ha un factor masculí lleu o moderat.

- **Fecundació *in vitro* amb injecció intracitoplasmàtica (FIV-ICSI):** Aquesta tècnica és la considerada més invasiva i també es realitza la fecundació de forma externa. En aquest cas, l'obtenció dels oòcits és igual a la FIV convencional, però, les cèl·lules del cúmulus del voltant de l'oòcit són eliminades. Un cop preparada i capacitada la mostra de semen, només s'agafa un sol espermatozoide, el qual se li eliminarà per escissió la cua i el cap es microinjectarà directament dins el citoplasma de l'oòcit (Figura 1.12C). Per aquesta tècnica, només és necessari un espermatozoide per cada oòcit a fertilitzar, per tant, és recomanada quan existeix un factor masculí sever com oligozoospèrmia o criptozoospèrmia o quan és necessari realitzar una biòpsia testicular per obtenir espermatozoides. Tot i així, actualment, quasi la meitat de tots els cicles de TRA utilitzen la FIV-ICSI (de Geyter *et al.*, 2018), ja que s'utilitza encara que no hi hagi recomanació específica, fins i tot en parelles amb infertilitat d'origen desconegut, degut a les elevades taxes de fertilització i embaràs (de Geyter *et al.*, 2018).

L'ús de les tècniques invasives com són la FIV convencional i la FIV-ICSI també estan associades a certs riscs. L'ús d'aquestes tècniques presenta un increment de defectes congènits comparat amb embarassos naturals (Sala *et al.*, 2011; Pan *et al.*, 2018). També, s'ha determinat un augment de les alteracions cromosòmiques i de *l'imprinting* parental en els nadons nascuts per aquestes tècniques en comparació a la concepció natural (Devroey and Van Steirteghem, 2004; Bonduelle *et al.*, 2005; Maher, 2005; Pan *et al.*, 2018).

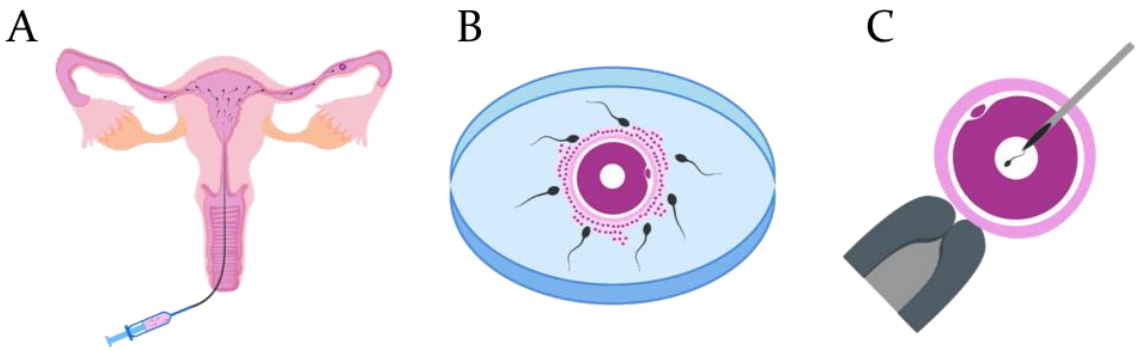


Figura 1.12. Tècniques de reproducció assistida. A) Inseminació intrauterina artificial (IIA). B) Fecundació *in vitro* (FIV) convencional. C) Fecundació in vitro amb injecció intracitoplasmàtica (FIV-ICSI). Adaptat de www.reproduccionasistida.org.

1.5.4. Infertilitat d'origen desconegut

Un dels reptes en el camp de la reproducció rau en l'assessorament de les TRAs en aquells pacients que aparentment tots els paràmetres es troben dins la normalitat però no aconsegueixen una gestació espontània. Aquest tipus de pacients infèrtils reben el nom d'individus normozoospèrmics amb infertilitat d'origen desconegut. Són pacients on el seminograma és aparentment normal (Veure apartat 1.5.2), és a dir, tots els paràmetres es troben dins el rang normal establert per la WHO (World Health Organization, 2010) i les proves complementàries no revelen ni defectes intrínsecs ni extrínsecs que expliquin la causa de la subfertilitat. Per tant, en aquest tipus de pacients és molt difícil aconsellar i/o predir quin tipus de tractament de reproducció assistida tindrà una taxa d'èxit més elevada i si és més adient utilitzar tècniques més invasives o tècniques menys invasives per aconseguir una gestació (Veure apartat 1.5.3), tenint en compte el cost econòmic i psicològic.

Clàssicament, el fracàs de les TRAs s'associava a defectes de l'oòcit, ja que de manera generalitzada es creia que l'oòcit aportava tots els elements necessaris per a l'embrió en desenvolupament i el paper de l'espermatozoide es basava solament en l'aportació del genoma haploide. Tot i així, es va observar en un estudi en què s'utilitzaven oòcits provinents de donació, que la causa de la infertilitat podia raure en l'individu masculí (Goudakou *et al.*, 2012). En els individus normozoospèrmics amb infertilitat d'origen desconegut segurament l'etiologia de la subfertilitat és post-espermatogènica, és a dir, que les causes poden trobar-se en la maduració epididimària en què l'espermatozoide no hagi adquirit el potencial de fertilització, en la capacitació i reacció acrosòmica, en la fertilització o bé condicionant el desenvolupament embrionari (Bracke *et al.*, 2018). Totes aquestes possibles causes no són detectades amb la realització d'un seminograma i és necessari l'ús de tècniques més sofisticades per poder determinar les causes subjacents en aquests tipus de pacients i, en un futur, aconsellar sobre el tipus de TRAs més adients. Tot i que s'han realitzat molts estudis tan descriptius com comparatius de diferents fenotips infèrtils, fins a dia d'avui, pocs estudis s'han focalitzat en els individus normozoospèrmics amb infertilitat d'origen desconegut (Garrido *et al.*, 2009; García-Herrero *et al.*, 2010a; Abu-Halima *et al.*, 2013; Bansal *et al.*, 2015; Salas-Huetos *et al.*, 2016; Selvam *et al.*, 2019).

1.6. Tècniques d'alt rendiment per a l'estudi molecular de l'espermatozoide

Les tècniques d'alt rendiment o –òmiques, es defineixen com l'estudi de la totalitat de molècules d'un tipus cel·lular, líquid o teixit. En els últims anys, arran de la reducció dels seus costs, els avenços en les tècniques i la seva posterior anàlisi han provocat un augment en el seu ús aplicat a diferents disciplines (Ritchie *et al.*, 2015). Breument, les tècniques més àmpliament utilitzades són la genòmica i l'epigenòmica, la transcriptòmica i la proteòmica (Figura 1.13).

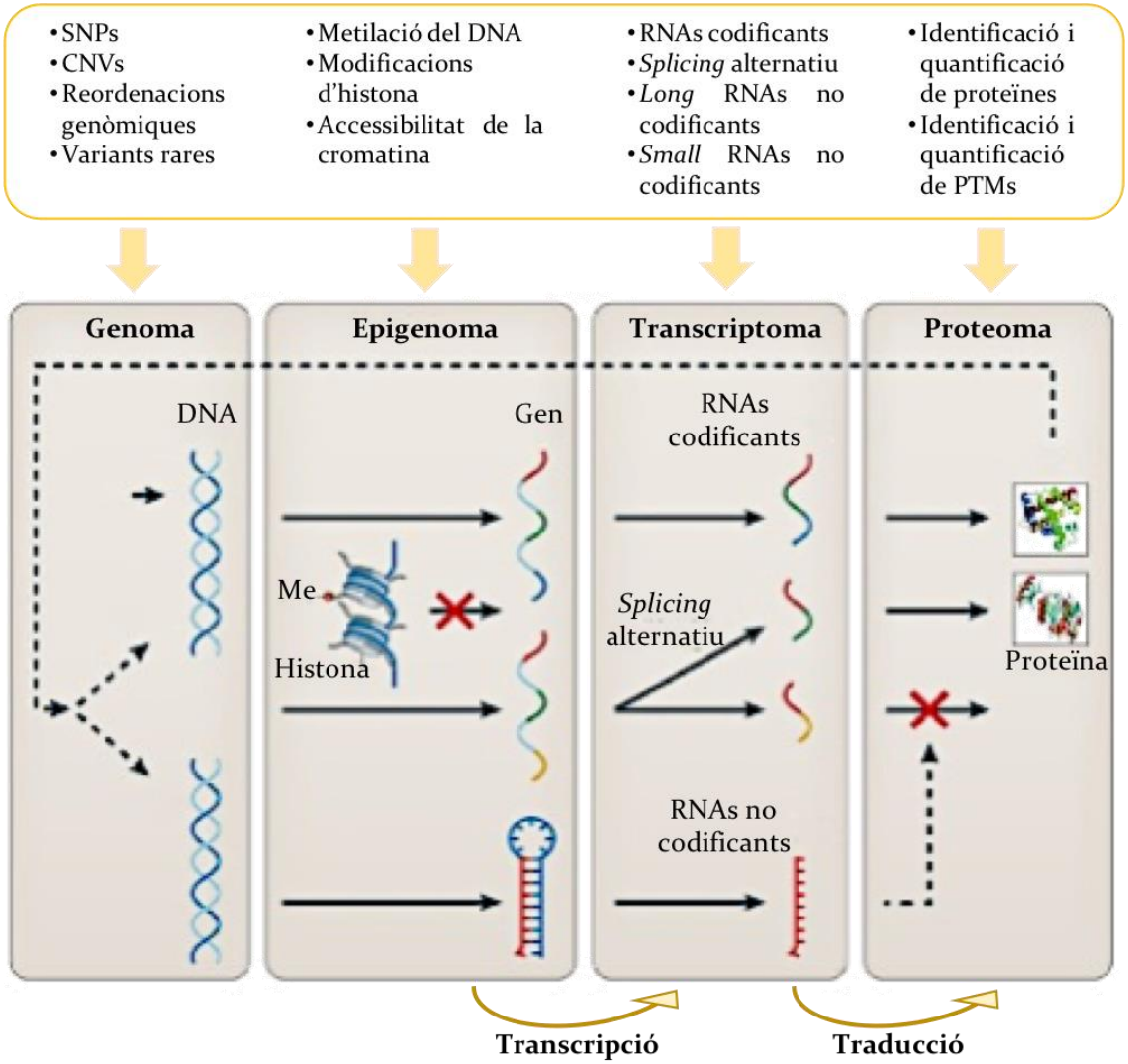


Figura 1.13. Tècniques d'alt rendiment. Les tècniques d'alt rendiment permeten estudiar el genoma i l'epigenoma, el transcriptoma i el proteoma, entre altres. CNV: *Copy number variation*; Me: Metilació; PTM: Modificació posttraduccional; SNP: *single-nucleotide polymoprhism*. Adaptat de (Ritchie *et al.*, 2015).

L'aplicació de tècniques d'alt rendiment a l'estudi de l'espermatozoide humà ha permès obtenir grans avenços en la coneixença de la seva composició i funció normal tot i la gran heterogeneïtat tan interindividual com intraindividual que presenten les mostres de semen (Krawetz *et al.*, 2011; Baker *et al.*, 2013; Wang *et al.*, 2013a; Jodar *et al.*, 2013, 2016b; Sendler *et al.*, 2013; Amaral *et al.*, 2014a; Castillo *et al.*, 2015; Pantano *et al.*, 2015; Carrell *et al.*, 2016; Manzoni *et al.*, 2018). També s'ha de tenir en compte que, ja que la cèl·lula espermàtica no transcriu ni tradueix, durant la maduració posttesticular, la capacitació i la reacció acrosòmica (Veure apartat 1.2), l'espermatozoide adquireix noves proteïnes i RNAs i les molècules contingudes pateixen canvis a nivell posttraduccional (Castillo *et al.*, 2018, 2019; Jodar, 2019).

Les tècniques d'alt rendiment han estat àmpliament utilitzades per a l'estudi de possibles causes subjacents de la infertilitat masculina associada a alteracions dels paràmetres seminals que reflecteixen desregulacions en l'espermatogènesi (Veure annex 2). Tot i així, encara manquen estudis per determinar l'impacte de la participació paterna a l'embrió en desenvolupament. En aquest sentit, principalment, mitjançant l'ús de tècniques d'alt rendiment es poden diferenciar dos tipus d'aproximacions: els estudis descriptius amb inferència *in silico* de les possibles funcions en la fertilització i desenvolupament embrionari i, els estudis comparatius entre diferents fenotips de pacients amb infertilitat d'origen desconegut. En els estudis comparatius, la comparació d'espermatozoides provinents d'individus normozoospèrmics amb la taxa d'èxit de diferents tipus de TRAs sense un factor femení conegut, permet caracteritzar i diferenciar la contribució espermàtica en la fertilització i les diferents etapes del desenvolupament embrionari. Concretament, la comparació d'individus normozoospèrmics que tenen o no èxit després de la IIA dona informació important en possibles defectes de l'espermatozoide que impedeixen que arribi a fecundar l'oòcit a l'oviducte. En canvi, la comparació d'individus normozoospèrmics que tenen o no èxit després de l'ús de FIV convencional pot donar explicació a possibles alteracions en molècules implicades en el reconeixement, interacció, fusió i penetració de l'espermatozoide en la fertilització de l'oòcit, així com del desenvolupament embrionari. De manera semblant, la comparació d'individus normozoospèrmics que tenen o no èxit després de l'ús de FIV-ICSI pot explicar alguns tipus de fallides de fertilització que tenen lloc després de l'entrada de l'espermatozoide a l'oòcit, així com la participació de l'espermatozoide a l'embriogènesi. Particularment, l'avaluació de la qualitat embrionària després de l'ús de FIV-ICSI d'espermatozoides provinents d'individus normozoospèrmics pot donar informació de la contribució paterna a l'embrió primerenc. Totes aquestes aproximacions mitjançant tècniques d'alt rendiment també contribueixen en la identificació de biomarcadors pronòstics que podrien predir la taxa d'èxit de les TRAs. En la present Tesi Doctoral ens centrarem només en les tècniques d'alt rendiment a nivell d'RNA (transcriptòmica) i de proteïnes (proteòmica).

1.6.1. Transcriptòmica

La transcriptòmica es defineix com a l'estudi de totes les molècules d'RNA presents en una cèl·lula o un conjunt de cèl·lules en un moment determinat. El terme transcriptoma engloba tan els RNA codificants que són traduïts a proteïna com els RNAs no codificants que no es tradueixen i regulen l'expressió gènica (Figura 1.13) (Ritchie *et al.*, 2015). Els primers estudis transcriptòmics es basaven en *microarrays* i presentaven una capacitat limitada d'identificar i quantificar la diversitat d'RNAs expressats en diferents tipus cel·lulars i en diferents genomes, principalment perquè no permetien el descobriment de noves isoformes o RNAs i requerien mètodes sofisticats de normalització (Wang *et al.*, 2009; Ozsolak and Milos, 2011). El desenvolupament de les tècniques de seqüenciació massiva *next-generation sequencing* (NGS) van permetre l'anàlisi d'RNAs a través de la seva conversió a cDNA i posterior seqüenciació a gran escala, creant-se així l'RNA-seq (Ozsolak and Milos, 2011). Així doncs, l'RNA-seq proporciona una mesura precisa dels nivells de tots els RNAs i de les seves isoformes en comparació a altres mètodes (Wang *et al.*, 2009).

Per a realitzar estudis d'RNA-seq, primerament es purifica una població d'RNA, que pot ser total o fraccionat (per exemple seleccionant els RNAs poliadenilats (poli $(A)^+$) o eliminant els RNAs ribosomals (rRNAs)). Aquests RNAs per retrotranscripció són passats a cDNA on se'ls uneixen uns adaptadors a un o ambdós extrems que permeten la seqüenciació (Figura 1.14). Seguidament, es pot realitzar una amplificació del material si es parteix de molt poca quantitat i a continuació se seqüencien totes les molècules per obtenir seqüències curtes (*reads*) d'un extrem (*single-end*) o dels dos extrems (*paired-end*) (Figura 1.14). Els *reads* solen ser de 30 a 400 bp, dependent de la tecnologia utilitzada. Després de la seqüenciació massiva, els *reads* obtinguts s'alineen al genoma de referència i es poden analitzar les dades obtingudes (Figura 1.14) (Wang *et al.*, 2009).

L'espermatozoide és una cèl·lula que no transcriu ni tradueix, per tant, quan s'estudia la cèl·lula espermàtica madura reflecteix el *cargo* que aportarà en el moment de la fertilització. Pel que fa a l'estudi de la cèl·lula espermàtica mitjançant RNA-seq s'han de tenir en consideració diferents aspectes degut a les seves particularitats. S'estima que la quantitat d'RNA present en un espermatozoide madur és unes 600 vegades inferior al d'una cèl·lula somàtica (Zhao *et al.*, 2006) que, en l'espermatozoide humà és de 10 – 50 fg per cèl·lula, ≈ 50 fg de *long* RNAs (> 200 nt) i 0.3 fg de *small* ncRNAs (< 200 nt) (Goodrich *et al.*, 2013; Jodar *et al.*, 2013). Això fa que petites contaminacions de l'ejaculat amb altres cèl·lules somàtiques, com ara leucòcits, cèl·lules epitelials i cèl·lules germinals immadures puguin emmascarar els resultats de l'RNA-seq d'espermatozoide (Jodar *et al.*, 2015, 2016b; Jodar, 2019). Aquestes cèl·lules contaminants, doncs, s'han d'eliminar mitjançant mètodes de purificació com la separació per gradients de densitat de sílice col·loïdal o la tècnica de *swim-up* que permet seleccionar els espermatozoides de més elevada motilitat

(Goodrich *et al.*, 2013; Mao *et al.*, 2013; Jodar *et al.*, 2015). Per purificar cèl·lules somàtiques també s'ha utilitzat el tampó de lisis cel·lular somàtica (*somatic cell lysis buffer*) però, tot i que és efectiu eliminant aquests contaminants, causen dany a la membrana de l'espermatozoide, provocant la pèrdua de seqüències mitocondrials i augmentant el risc de pèrdua de transcrits de la peça intermèdia (Mao *et al.*, 2013). Un cop els espermatozoides estan purificats, també s'ha d'assegurar la seva completa lisi per tal d'alliberar tots els RNAs, majoritàriament continguts dins el nucli (Johnson *et al.*, 2015). Per dur-ho a terme i degut a l'elevada quantitat de ponts disulfur que contenen les protamines que empaqueten el DNA de l'espermatozoide, l'addició d'agents reductors com el beta-mercaptoetanol al tampó de lisi és essencial per a l'alliberació dels RNAs i posterior extracció (Goodrich *et al.*, 2013).

Degut a les característiques específiques de l'RNA de l'espermatozoide que majoritàriament es troba fragmentat, no es poden utilitzar tècniques per a l'avaluació de la integritat, com l'ús estandarditzat del valor de l'*RNA integrity number* (RIN) en cèl·lules somàtiques (Schroeder *et al.*, 2006). Tot i així, se solen aplicar altres controls de qualitat que s'exposen a continuació. Posterior a l'extracció de l'RNA d'espermatozoide, l'RNA obtingut es tracta amb DNases per eliminar la possible contaminació de DNA. Al contrari que el DNA, l'RNA no conté introns, de manera que l'absència de contaminació de DNA es verifica mitjançant l'amplificació directa de l'RNA amb *primers* específics d'un intró o que continguin un intró a través de PCR a temps real. Si no hi ha contaminació de DNA, no hi haurà amplificació, mentre que si hi ha amplificació indica la presència de DNA i per tant, contaminació. La integritat de l'RNA de l'espermatozoide i l'absència de contaminants somàtics s'avalua mitjançant la retrotranscripció de l'RNA a cDNA seguit de PCR a temps real, normalment dels gens continguts específicament de l'espermatozoide *PRM1* o *PRM2* i del marcador específic de leucòcits *receptor-type tyrosine-protein phosphatase C* (*PTPRC)* expressat principalment en cèl·lules somàtiques, respectivament (Goodrich *et al.*, 2013; Jodar *et al.*, 2015, 2016b; Gòdia *et al.*, 2018).

Per a dur a terme experiments d'RNA-seq en altres tipus cel·lulars, normalment es requereix fins a 1 µg d'RNA, però en l'espermatozoide la quantitat inicial sol ser d'uns 2 – 10 ng, que implica l'ús d'un mínim de 2 milions d'espermatozoides (Goodrich *et al.*, 2013; Jodar *et al.*, 2013, 2015; Johnson *et al.*, 2015; Gòdia *et al.*, 2018). També, una metodologia àmpliament utilitzada en altres tipus cel·lulars és l'enriquiment d'RNAs codificants a través de la selecció per poli $(A)^+$, però aquest tipus de selecció amb la cèl·lula espermàtica s'ha demostrat que pot comportar un biaix a l'extrem 3', degut a la no selecció d'RNAs sense cua de poli (A) i a l'abundància d'RNAs fragmentats, entre altres (Gòdia *et al.*, 2018). Així doncs, l'estratègia més utilitzada en espermatozoide madur és l'RNA-seq de l'RNA total, tot i que es perden fraccions específiques de *small* ncRNAs (< 200 nt) que requereixen protocols específics i

selecció per mida per a poder-los seqüenciar (Krawetz *et al.*, 2011; Pantano *et al.*, 2015).

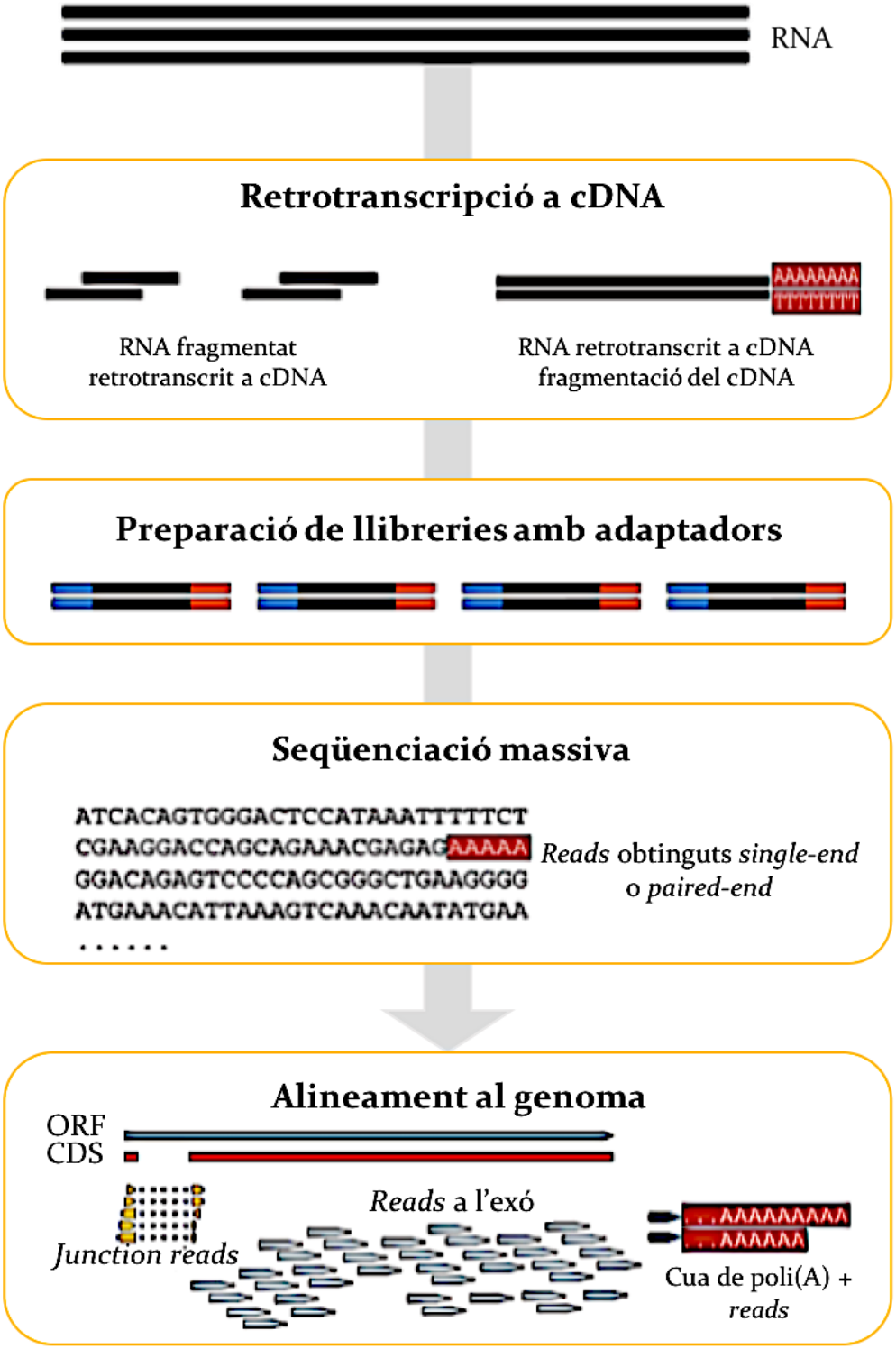


Figura 1.14. Metodologia típica d'un estudi d'RNA-seq. Breument, els RNAs primer són fragmentats i convertits a cDNA o bé primer convertits a cDNA i seguidament fragmentats*; seguidament es preparen les llibreries i s'afegeixen els adaptadors i una seqüència és obtinguda per cada fragment de cDNA utilitzant seqüenciació massiva. Finalment, els *reads* s'alineen al genoma. *En espermatozoide no és necessària la fragmentació dels RNAs, ja que l'RNA de la cèl·lula espermàtica es troba altament fragmentat. Adaptat de (Wang *et al.*, 2009).

Utilitzant RNA-seq amb RNA total es troba un enriquiment en rRNAs fragmentats i RNAs mitocondrials (mitoRNAs) (Jodar *et al.*, 2013; Mao *et al.*, 2014). Per analitzar les dades d'RNA-seq de l'espermatozoide, un cop alineades al genoma de referència, s'ha utilitzat àmpliament la quantificació de l'abundància dels transcrits anotats, en la qual es consideren el número total de *reads* que alineen a un transcrit, després d'ajustar-ho per la llargada total del transcrit i, s'obtenen, o bé els fragments per quilobase de transcrit per milió (FPKM) quan es tracta d'experiments *paired-end,* o bé *reads* per quilobase de transcrit per milió (RPKM) quan es tracta d'experiments *single-end*.

1.6.1.1. Estudis descriptius amb inferència *in silico*

Gràcies a l'ús de les tècniques d'RNA-seq, en humà, s'ha descrit una complexa població d'RNAs tan codificants com no codificants (Taula 1.3), la majoria d'ells continguts al cap de l'espermatozoide (Krawetz *et al.*, 2011; Jodar *et al.*, 2013; Sendler *et al.*, 2013; Johnson *et al.*, 2015; Pantano *et al.*, 2015; Jodar, 2019). Concretament, s'ha establert que al voltant del 80 % d'RNAs de l'espermatozoide humà són rRNAs (Jodar *et al.*, 2013; Johnson *et al.*, 2015; Gòdia *et al.*, 2018; Jodar, 2019) i es troben selectivament fragmentats, segurament per assegurar que l'espermatozoide madur sigui traduccionalment inert (Johnson *et al.*, 2011; Jodar, 2019).

Malgrat que la majoria d'RNAs codificants de l'espermatozoide es troben fragmentats, n'hi ha una part que es troben intactes. Aquests RNAs intactes es troben relacionats amb la fertilitat i el desenvolupament embrionari (Sendler *et al.*, 2013) i es conserva en diferents espècies com humà, ratolí, rata i conill (Schuster *et al.*, 2016). Un dels mecanismes proposats per la funcionalitat dels RNAs codificants de l'espermatozoide en el desenvolupament embrionari és que aquests es conserven intactes fins el moment de la fertilització i, un cop alliberats a l'oòcit, es tradueixen mitjançant la maquinària traduccional activa de l'oòcit fecundat (Jodar *et al.*, 2013; Ntostis *et al.*, 2017; Gòdia *et al.*, 2018; Jodar, 2019). En ratolí, mitjançant RNA-seq i inferència *in silico*, s'ha proposat que alguns RNAs codificants intactes de l'espermatozoide es tradueixen i col·laboren en l'eliminació dels RNAs materns que té lloc abans de l'activació del genoma embrionari i estableixen un punt de control per a l'inici de l'activació del zigot (Ntostis *et al.*, 2017).

A més d'RNAs codificants, l'espermatozoide madur conté una població única i complexa d'RNAs no codificants tan *long* ncRNAs com *small* ncRNAs (Krawetz *et al.*, 2011; Jodar *et al.*, 2013; Pantano *et al.*, 2015). Dins els *small* ncRNAs trobem fragments de RNA transferència (tRFs), miRNAs, *small nucleolar* RNAs (snoRNAs), *small nuclear* RNAs (snRNAs), RNAs derivats de Y (YRNAs), siRNAs, piRNAs i *small* RNAs derivats del mitocondri (mtRNAs) (Jodar *et al.*, 2013). Tanmateix, alguns RNAs no codificants sembla que són exclusius de l'espermatozoide i no s'han identificat

en altres tipus cel·lulars (Jodar *et al.*, 2013; Sendler *et al.*, 2013; Gòdia *et al.*, 2018; Jodar, 2019). N'és un clar exemple que un elevat nombre de transcrits de l'espermatozoide madur provenen de regions intròniques que són retingudes específicament i reben el nom d'elements intrònics retinguts (IEs), ja que no s'observen en testicle i només s'observen a l'espermatozoide madur. Tot i així, la seva funcionalitat encara no ha estat descrita (Jodar *et al.*, 2013).

1.6.1.2. Estudis comparatius entre fenotips

La determinació d'alteracions en el perfil d'RNAs en humà ha estat àmpliament utilitzat per a l'estudi de possibles causes de la infertilitat masculina associada a alteracions dels paràmetres seminals (Taula 1.3) (Platts *et al.*, 2007; Garrido *et al.*, 2009; García-Herrero *et al.*, 2010a; Jodar *et al.*, 2012; Montjean *et al.*, 2012; Abu-Halima *et al.*, 2013; Bansal *et al.*, 2015; Salas-Huetos *et al.*, 2016, 2015; Zhang *et al.*, 2019a; Heidary *et al.*, 2019; Caballero-Campo *et al.*, 2020). També, alguns estudis han predit, mitjançant *microarray* o tècniques d'RNA-seq, tan per RNAs codificants com per RNAs no codificants, la taxa d'èxit de diferents TRAs com serien la IIA o la FIV-ICSI (Taula 1.3) (García-Herrero *et al.*, 2010b, 2010a, 2011; Bonache *et al.*, 2012; Jodar *et al.*, 2015; Burl *et al.*, 2018; Xu *et al.*, 2020).

Tot i així, com s'ha mencionat anteriorment, molts RNAs de l'espermatozoide es troben fragmentats i les anàlisis tradicionals a nivell de transcrit no són útils en estudis d'RNA-seq d'espermatozoide on es comparen diferents fenotips. Per superar aquesta limitació, es va establir el que s'anomenen els *sperm RNA elements* (SREs) (Sendler *et al.*, 2013; Jodar *et al.*, 2015) que són seqüències curtes que corresponen a exons, parts intròniques o intergèniques que tenen cert número de *reads* i s'alineen independentment de la seva anotació encara que siguin contigües (Sendler *et al.*, 2013; Jodar *et al.*, 2015; Johnson *et al.*, 2015; Gòdia *et al.*, 2018; Estill *et al.*, 2019; Swanson *et al.*, 2020). Per tant, aquesta metodologia és idònia per estudis comparatius i permet l'anàlisi sense biaix de diferents perfils d'RNAs. N'és un exemple l'anàlisi de SREs que va permetre la identificació de 648 SREs que són requerits per la concepció en individus normozoospèrmics i que la seva presència estaria relacionada amb millor taxa d'èxit en la utilització de TRAs menys invasives com la IIA. L'absència d'almenys un d'aquests elements es correlacionava en una baixa taxa d'èxit de la IIA (Jodar *et al.*, 2015).

Taula 1.3. Estudis del contingut d'RNAs tan codificants com no codificants de l'espermatozoide humà madur. S'han incorporat estudis descriptius i comparatius segons els paràmetres seminals i la potencial contribució a l'embrió. S'han inclòs estudis de *microarray* i RNA-seq.

Contingut d'RNAs	Tipus RNAs	Estudi
Descripció d'RNAs		
Compilació dels RNAs de l'espermatozoide	*RNAs codificants*	(Ostermeier *et al.*, 2002); (Ostermeier *et al.*, 2004); (Jodar *et al.*, 2013); (Sendler *et al.*, 2013); (Johnson *et al.*, 2015); (Schuster *et al.*, 2016)
	RNAs no codificants	(Krawetz *et al.*, 2011); (Jodar *et al.*, 2013); (Pantano *et al.*, 2015)
Comparativa perfil d'RNAs amb paràmetres seminals		
Normozoospèrmia amb infertilitat d'origen desconegut	*RNAs codificants*	(Garrido *et al.*, 2009); (García-Herrero *et al.*, 2010a); (Bansal *et al.*, 2015)
	RNAs no codificants	(Abu-Halima *et al.*, 2013); (Salas-Huetos *et al.*, 2016)
Astenozoospèrmia	*RNAs codificants*	(Jodar *et al.*, 2012); (Bansal *et al.*, 2015); (Caballero-Campo *et al.*, 2020)
	RNAs no codificants	(Abu-Halima *et al.*, 2013); (Salas-Huetos *et al.*, 2015); (Bansal *et al.*, 2015); (Zhang *et al.*, 2019a); (Heidary *et al.*, 2019)
Oligozoospèrmia	*RNAs codificants*	(Montjean *et al.*, 2012)
	RNAs no codificants	(Abu-Halima *et al.*, 2013); (Salas-Huetos *et al.*, 2015)
Teratozoospèrmia	*RNAs codificants*	(Platts *et al.*, 2007)
	RNAs no codificants	(Salas-Huetos *et al.*, 2015)
Comparativa del perfil d'RNAs per determinar la contribució de l'espermatozoide a l'embrió		
Qualitat embrionària i èxit de les TRAs	*RNAs codificants*	(García-Herrero *et al.*, 2010a); (García-Herrero *et al.*, 2010b); (García-Herrero *et al.*, 2011); (Bonache *et al.*, 2012); (Jodar *et al.*, 2015)
	RNAs no codificants	(Jodar *et al.*, 2015); (Xu *et al.*, 2020)

TRA: Tècnica de Reproducció Assistida

1.6.2. Proteòmica

La proteòmica és l'estudi de totes les proteïnes que conformen una cèl·lula, fluid o teixit d'un organisme en un determinat moment i sota unes condicions concretes (el que es coneix com el proteoma) fent servir principalment tècniques basades en espectrometria de masses (MS) (Aebersold and Mann, 2003). Les tècniques d'anàlisi proteòmica permeten identificar i quantificar tan les proteïnes i les seves isoformes com les PTMs associades (Figura 1.13) (Ritchie *et al.*, 2015).

Pel que fa a l'anàlisi de l'espermatozoide, tot i que sigui una cèl·lula inactiva a nivell de transcripció i traducció, durant la seva maduració i els processos de capacitació i reacció acrosòmica perd proteïnes, n'adquireix de noves i experimenta canvis a nivell de PTMs, en especial la fosforilació de residus tirosina, la qual cosa implica canvis en el seu proteoma (Castillo *et al.*, 2019). Degut a les particularitats de l'espermatozoide, en condicions normals, són requerits almenys 5 milions d'espermatozoides per a realitzar estudis de proteòmica (Veure annex 2). A més, és necessari que la mostra tingui la màxima puresa cel·lular possible i estigui lliure de contaminants com serien proteïnes del plasma seminal, cèl·lules somàtiques com els leucòcits i cèl·lules epitelials i cèl·lules germinals immadures (Veure annex 2). Per assolir aquesta puresa, normalment es realitza una separació per gradients de densitat de sílice col·loïdal que permet l'obtenció d'una població d'espermatozoides òptima. Si es vol aïllar les cèl·lules amb la motilitat més elevada es pot realitzar un *swim-up* (Codina *et al.*, 2015). En cas de necessitar i assegurar la completa eliminació de leucòcits, es pot procedir a fer una selecció per boles magnètiques del marcador específic PTPRC (Amaral *et al.*, 2014a; Codina *et al.*, 2015; Kramer, 2016). Seguidament, els espermatozoides són lisats i solubilitzats (Amaral *et al.*, 2014a; Codina *et al.*, 2015).

Les tècniques tradicionals de proteòmica, es basaven en la separació de les proteïnes, seguit del seu aïllament, digestió mitjançant proteases i la posterior identificació i anàlisi dels pèptids obtinguts. Les tècniques basades en l'electroforesi en gels de poliacrilamida en presència de SDS (*sodium dodecyl sulfate polyacrylamide gel electrophoresis*; SDS – PAGE), han estat àmpliament utilitzades. Es pot realitzar la primera dimensió d'una electroforesi bidimensional (2D), separant per pes molecular i, després fent una segona dimensió o separació mitjançant enfocament isoelèctric, el qual separa les proteïnes pel seu punt isoelèctric, resultant en una separació dels punts proteics (*protein spots*). El següent pas és l'escissió dels punts proteics i de les bandes proteiques, la seva elució i digestió a pèptids en el gel (*in-gel digestion*) i, finalment, els pèptids generats s'identifiquen directament per MS (Figura 1.15; veure annex 2). Alternativament, també es pot fer només una electroforesi monodimensional (1D) de les proteïnes, separant solament per pes molecular el conjunt de bandes electroforètiques, però en aquest cas cal eluir les proteïnes, procedir a la seva digestió a pèptids i a la separació mitjançant

cromatografia líquida (LC) monodimensional seguit de la identificació per MS (Figura 1.15; veure annex 2).

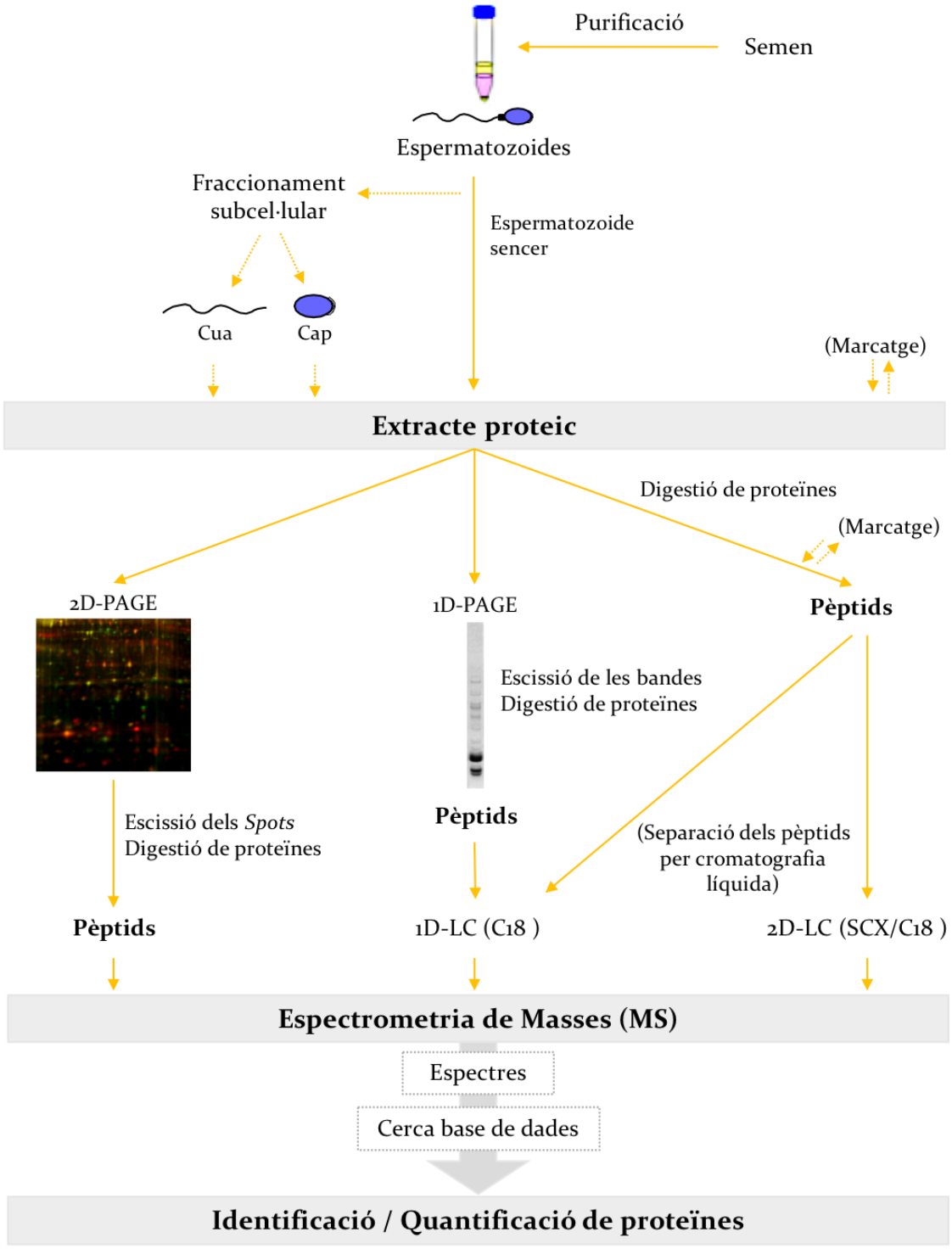


Figura 1.15. Proteòmica de l'espermatozoide. Els espermatozoides es purifiquen i es pot estudiar l'espermatozoide sencer o realitzar un subfraccionament cel·lular. S'obtenen els extractes proteics i seguidament es pot realitzar un gel bidimensional o monodimensional, escindir els punts proteics (*spots*) o bandes, respectivament, digerir les proteïnes i analitzar-ho per espectrometria de masses (part esquerra i central) o bé digerir les proteïnes a pèptids sense separació i analitzar-ho, també, per espectrometria de masses (part dreta). Finalment els espectres se cerquen a una base de dades que permet la identificació i quantificació de les proteïnes. Adaptat de l'Annex 2 (Jodar *et al.*, 2017).

Més recentment i degut als avanços en la sensibilitat de les màquines, s'ha utilitzat àmpliament la digestió directa de les proteïnes (sense cap tipus de pre-separació) per tal de generar pèptids (*in-solution digestion*), normalment fent servir la proteasa tripsina i separació per LC acoblada a MS en tàndem (LC-MS/MS; Figura 1.15; veure annex 2). En aquesta metodologia es realitza la identificació en el MS1 i la fragmentació en el MS2 i s'obté la intensitat en relació a la massa / càrrega (m / z) dels ions que dona lloc als espectres. Els espectres són contrastats en bases de dades i permeten la identificació i quantificació dels pèptids i, per tant, de les proteïnes (Figura 1.15; veure annex 2). Una altra opció que s'ha aplicat més recentment és la doble separació dels pèptids per LC abans de la identificació per MS/MS, que es coneix com a 2D-LC-MS/MS (Figura 1.15).

1.6.2.1. Estudis descriptius amb inferència *in silico*

L'aplicació de tècniques d'alt rendiment per estudiar el perfil proteic de l'espermatozoide humà han permès incrementar el coneixement actual del gàmeta masculí, la seva composició i la seva funcionalitat (Taula 1.4). La primera compilació del proteoma de l'espermatozoide va permetre la identificació de 6198 proteïnes no redundants (Amaral *et al.*, 2014a). L'any 2017 es va actualitzar el total de les proteïnes de l'espermatozoide i se'n van compilar un total de 6238, de les quals 1430 corresponien a la fracció del cap i 1516 a la fracció de la cua (Veure annex 2). Més recentment, s'han compilat un total de 6871 proteïnes en la cèl·lula espermàtica madura les quals, en el moment de la fertilització s'alliberen a l'oòcit (Castillo *et al.*, 2018).

Com s'ha mencionat anteriorment (Veure apartat 1.4.4), s'ha evidenciat molt el paper dels RNAs, sobretot els RNAs no codificants, al desenvolupament embrionari primerenc, però, tot i que l'espermatozoide aporta quasi 7000 proteïnes en el moment de la fertilització, el paper d'aquestes ha estat menys indagat. No obstant, s'han descrit un conjunt de proteïnes amb possibles funcions durant les primeres fases del desenvolupament embrionari (Amaral *et al.*, 2014a; Jodar *et al.*, 2016a; Castillo *et al.*, 2018). Concretament, s'han descrit proteïnes, en les quals s'hi troben factors de transcripció, zinc *fingers,* variants d'histona, modificadors d'histones i proteïnes relacionades amb el metabolisme o regulació del DNA i l'RNA que podrien ser crucials per l'expressió gènica tan a nivell transcripcional com traduccional en el desenvolupament embrionari (Castillo *et al.*, 2014a, 2014b, 2015, 2018). En l'estudi *in silico* de Castillo et al., 2018, es van identificar proteïnes de l'espermatozoide funcionalment relacionades amb termes d'ontologia gènica com són la fertilització, el desenvolupament embrionari primerenc i la regulació de l'expressió gènica (Castillo *et al.*, 2018). Aquesta anàlisi *in silico* va permetre identificar 103 proteïnes implicades en el procés de fertilització, com ara les proteïnes IZUMO1 i PLCζ1 i 93 proteïnes relacionades amb el desenvolupament embrionari primerenc, des de la

formació del zigot fins al blastocist, en 59 de les quals el model KO en ratolí presentava alteracions importants en l'embriogènesi. D'aquestes 59, 11 proteïnes es trobaven associades a letalitat embrionària en la fase de zigot, remarcant la importància de l'espermatozoide en les primeres etapes del desenvolupament embrionari abans de l'activació del genoma embrionari. També, 29 proteïnes es trobaven associades a alteracions en l'estadi de mòrula i 19 en l'estadi de blastocist (Castillo *et al.*, 2018). A més, es van identificar 560 proteïnes implicades en la regulació de l'expressió gènica en altres tipus cel·lulars: 381 factors de transcripció o proteïnes relacionades amb els factors de transcripció; 25 modificadors de la cromatina associats a la metilació del DNA; 118 modificadors de la cromatina associats a les PTMs d'histones i 36 proteïnes que regulen la transcripció, el processament i la funció dels RNAs no codificants (Castillo *et al.*, 2018). De les 560 proteïnes implicades en la regulació de l'expressió gènica, es va determinar que 28 podrien tenir una funció durant les primeres etapes del desenvolupament embrionari (Castillo *et al.*, 2018). Així doncs, la contribució proteica de l'espermatozoide a l'embrió, sembla ser que tingui un rol important. A més a més, mitjançant l'estudi integratiu de les dades transcriptòmiques i proteòmiques de l'espermatozoide, l'oòcit i l'embrió primerenc es van predir que algunes de les proteïnes del blastocist podrien tenir un origen patern, on s'inclouen factors de transcripció, modificadors d'histones, reguladors de l'activitat guanosina trifosfatasa (GTPasa) i proteïnes relacionades amb el centrosoma (Castillo *et al.*, 2018; Jodar, 2019). Concretament, aquesta anàlisi va predir que 108 proteïnes del blastocist podrien provenir de l'espermatozoide, bé per la seva detecció exclusiva en el proteoma de la cèl·lula espermàtica i la seva absència a nivell d'RNA de l'espermatozoide, oòcit i els diferents estadis del desenvolupament embrionari (82 proteïnes); o bé, per la presència exclusiva del transcrit a l'espermatozoide i l'absència a nivell de proteïna tan en l'espermatozoide i l'oòcit com en els fluids i les cèl·lules associades reproductores (26 proteïnes). Aquestes 26 proteïnes identificades en l'estadi de blastocist que es troben a nivell d'RNA a l'espermatozoide, podrien ser el producte de la traducció d'RNAs aportats exclusivament per l'espermatozoide i traduïts per la maquinària traduccional materna. Una d'aquestes potencials proteïnes traduïdes seria la DGCR8 que es troba implicada en el processament dels pri-miRNAs a miRNAs (Wang *et al.*, 2007; Castillo *et al.*, 2018). Una altra seria la *protein TALPID3* (KIAA0586), que la seva absència en ratolí dona lloc a embrions amb defectes en el conducte neural i en l'eix esquerra – dreta (Bangs *et al.*, 2011; Jodar, 2019).

1.6.2.2. Estudis comparatius entre fenotips

Amb les millores en les tècniques proteòmiques, ha crescut l'interès en poder comparar de manera replicable diferents fenotips a través de les estratègies *bottom-up* MS. Actualment, dues estratègies de quantificació són les més utilitzades, la quantificació sense marcatge (*label-free*) de les proteïnes o pèptids o el marcatge diferencial de les proteïnes o pèptids seguit de la seva identificació i quantificació. La quantificació proteòmica en la tècnica *label-free* es basa en la correlació directa entra la intensitat de la senyal MS dels pèptids amb la quantitat de proteïna relativa o absoluta o utilitzant el nombre d'espectres obtinguts per un determinat pèptid o proteïna com a indicador de la seva quantitat (Bantscheff *et al.*, 2007). Per contra, el marcatge diferencial es basa en la utilització d'isòtops estables per tal d'afegir una marca (*tag*) específica que pugui ser reconeguda per l'espectròmetre de masses i permeti la quantificació dels pèptids o proteïnes. Aquests *tags* poden unir-se als pèptids o proteïnes, fent servir estratègies metabòliques, químiques o enzimàtiques (Bantscheff *et al.*, 2007). Un dels marcatges diferencials més utilitzat són els *tags* isobàrics, dels quals n'és un exemple el *tandem mass tag* (TMT™), que es troba disponible comercialment (Pappireddi *et al.*, 2019). Les mostres es marquen cadascuna amb un *tag* i posteriorment són multiplexades en quantitats iguals, juntament amb un control intern format per alíquotes equivalents de cada mostra individual ajuntades en un únic control que permet la seva comparació entre elles. Els *tags* isobàrics són reactius que modifiquen covalentment pèptids i són afegits usualment després de la digestió de les proteïnes en pèptids. Presenten una massa total idèntica entre els diferents *tags,* però presenten diferències entre els isòtops (Figura 1.16). D'aquesta manera durant la fragmentació dels pèptids en el MS2, cada *tag* presenta un punt de fragmentació diferent amb una massa relativa diferent i permet discernir de quina mostra prové el pèptid. Els mateixos pèptids que provenen de diferents mostres, elueixen al mateix temps i presenten 1 sol espectre d'identificació al MS1, la qual cosa no incrementa la complexitat d'anàlisi, ja que serà durant la fragmentació del MS2 quan es podrà diferenciar l'origen del pèptid (Pappireddi *et al.*, 2019). L'energia que s'afegeix per a la fragmentació, només permet el trencament d'un enllaç que sol ser el punt de trencament previst per al *tag* isobàric. En trencar-se, el *tag* produeix *reporter ions* de baixa m / z que tenen diferent massa i són utilitzats per a la quantificació relativa (Figura 1.16) (Pappireddi *et al.*, 2019). Així doncs, les mostres marcades amb *tags* han esdevingut una metodologia molt utilitzada, ja que permet identificar i quantificar els pèptids (i per tant, les proteïnes) de diferents individus a un cost raonable a través de les estratègies de multiplexat. També permet, en cas de no detectar un pèptid en una mostra en concret, però sí en altres, deduir l'absència o minoria específica d'aquest pèptid específic en aquella mostra (Pappireddi *et al.*, 2019). Tot i la fiabilitat del marcatge diferencial, es requereix d'una anàlisi posterior estricta per a què els resultats obtinguts siguin d'una confiança elevada. Recentment, el nostre grup ha proposat

uns criteris de selecció estrictes per a la quantificació peptídica, en els quals els pèptids han de ser quantificats per almenys 2 *peptide spectrum matches* (PSMs), amb un coeficient de variació de menys del 50 % entre ells en almenys el 75 % de les mostres analitzades (descrit a (Barrachina *et al.*, 2019)).

La proteòmica comparativa amb o sense marcatge de l'espermatozoide madur d'individus normozoospèrmics amb diferents perfils infèrtils ha estat àmpliament explotada per determinar marcadors de (in)fertilitat (Taula 1.4). Entre ells, pacients que presenten alteracions en els paràmetres seminals, com ara alteracions en la concentració, motilitat o morfologia (Taula 1.4; veure annex 2). Aquestes comparacions a nivell proteòmic proporcionen una idea de possibles vies alterades que reflecteixen les alteracions dels paràmetres seminals i proporcionen informació valuosa dels mecanismes patològics que podrien ser la causa de la infertilitat. El perfil més àmpliament estudiat ha estat l'astenozoospèrmia on hi ha una alteració de la motilitat espermàtica (Zhao *et al.*, 2007; Martínez-Heredia *et al.*, 2008; Chan *et al.*, 2009; Siva *et al.*, 2010; Chao *et al.*, 2011; Parte *et al.*, 2012; Shen *et al.*, 2013; Amaral *et al.*, 2014b; Hashemitabar *et al.*, 2015; Liu *et al.*, 2015; Saraswat *et al.*, 2017; Nowicka-Bauer *et al.*, 2018; Sinha *et al.*, 2019; Guo *et al.*, 2019; Moscatelli *et al.*, 2019), segurament perquè l'obtenció d'una població pura d'espermatozoides és més difícil d'obtenir en fenotips oligozoospèrmics on hi ha menys quantitat d'espermatozoides (Botta *et al.*, 2009) i teratozoospèrmics on hi ha defectes morfològics (Liao *et al.*, 2009), ja que la purificació és més ineficient (Taula 1.4). Per altra banda, també s'han realitzat estudis proteòmics en pacients infèrtils amb causes conegudes d'infertilitat per aprofundir més en l'etiologia d'aquesta, com ara pacients amb varicocele o hipogonadisme-hipogonadotròpic (Arato *et al.*, 2020; Panner Selvam *et al.*, 2020).

En paral·lel a les tècniques clàssiques de *bottom-up* MS, han aparegut les tècniques *top-down* MS que es basen en l'estudi de les proteïnes intactes sense digestió prèvia (Smith and Kelleher, 2018). Aquesta estratègia és aplicable a proteïnes o pèptids relativament petits. En les tècniques de *bottom-up* MS la digestió peptídica disminueix la complexitat de la mostra en simplificar les molècules a analitzar i per tant, facilita la interpretació dels resultats, tot i que no es pot extrapolar informació de diferents modificacions d'una mateixa molècula. No obstant, en les tècniques de *top-down* MS, la complexitat de la mostra no es redueix i és necessari una anàlisi de dades complexa. La tècnica de *top-down* MS es basa en la identificació de la massa inicial de la proteïna intacta i la mesura de les seves proteoformes que informa sobre els efectes de diferents combinacions de PTMs, variacions de seqüència o proteïnes truncades que es produeixen en la mateixa molècula, donant-ne informació a nivell molecular. Aquests resultats permeten correlacionar diferents modificacions que tenen lloc en una mateixa molècula, la qual cosa permet connectar-ho directament amb aplicacions potencials en el diagnòstic (Kelleher *et al.*, 2014). Si una proteoforma truncada és la causant d'una alteració concreta, mitjançant tècniques

de *bottom-up* basades en la digestió peptídica no es podria identificar la proteïna truncada perquè es trobaria digerida, mentre que mitjançant *top-down* MS permetria connectar la proteoforma truncada a un diagnòstic concret. Tot i així, degut a la seva complexitat, aquesta tècnica, tal com s'ha comentat, se sol utilitzar només en l'estudi de proteïnes de baix pes molecular i, addicionalment es poden combinar les de tècniques *bottom-up* MS i *top-down* MS, ja que solen aportar informació complementària (Bonet-Costa *et al.*, 2012).

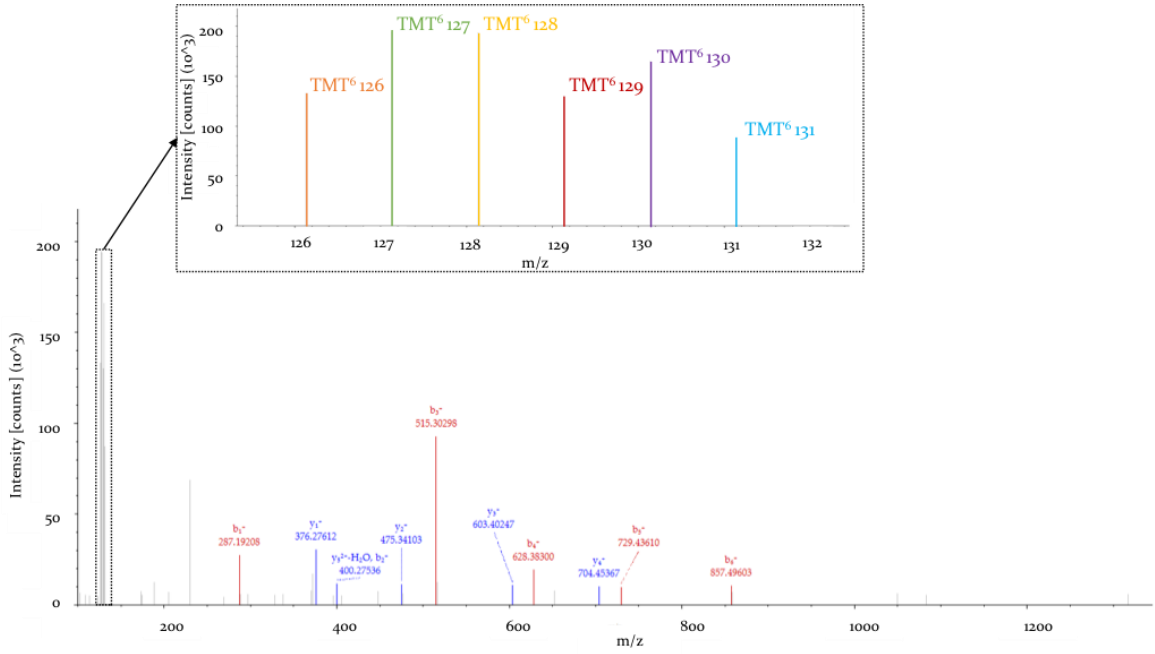


Figura 1.16 Exemple d'espectre obtingut al MS2 per la fragmentació d'un pèptid marcat amb TMT6-plex. Concretament, el pèptid correspon a la proteïna *dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex* (DLAT) (seqüència peptídica: GIDLTQVK, N-term-TMT6plex (229.16293 Da), K8-TMT6plex (229.16293 Da)). Al requadre s'observen els *ions reporters* (massa / càrrega (m / z) 126 – 131) on la intensitat de cada *ion reporter* representa l'abundància relativa del pèptid per cada mostra marcada. Adaptat de l'Annex 3 (Torra-Massana *et al.*, en preparació).

Com s'ha mencionat anteriorment, les protamines són les proteïnes més abundants del nucli de l'espermatozoide madur i empaqueten, en humans, aproximadament el 92 % del genoma patern (Veure apartat 1.4.3.2.2) (Gatewood *et al.*, 1987; Oliva, 2006; Jodar and Oliva, 2014). Tot i que s'han relacionat alteracions en el contingut de protamines en individus normozoospèrmics amb infertilitat d'origen desconegut a nivell proteic i defectes en el desenvolupament embrionari i la taxa d'èxit de les TRAs (Khara, 1997; Nasr-Esfahani *et al.*, 2004b, 2004a; Aoki *et al.*, 2005a, 2006b; de Mateo *et al.*, 2009; Simon *et al.*, 2011), aquests resultats es fonamenten en tècniques convencionals basades en l'extracció específica àcida de les protamines i la seva visualització a través de tècniques àcid – urea PAGE, del qual es calcula el ràtio

relatiu entre P1/P2 i pre-P2/P2. Per calcular l'abundància tan de P1 com dels membres de la família de P2 se sol incloure un estàndard de protamines de concentracions conegudes en el gel àcid – urea PAGE, del qual es realitza una recta de regressió i s'extrapola la quantitat de protamina de les mostres d'estudi (Mengual *et al.*, 2003). Tot i les àmplies evidències a nivell proteic que alteracions en el ràtio P1/P2 poden afectar el desenvolupament embrionari, un estudi va concloure que el ràtio P1/P2 en una població fèrtil, podia variar de 0,54 – 1,43 (Nanassy *et al.*, 2011), indicant que segurament es necessiten tècniques més sofisticades per establir realment el paper de les protamines en l'embriogènesi i la fertilitat, tot i que degut a les seves particularitats fisicoquímiques no es fàcil el seu estudi mitjançant espectrometria de masses (basicitat extremadament elevada, seqüència d'aminoàcids extremadament rica en residus arginina i cisteïna, mida petita i elevada similitud de seqüència entre les formes madures de P2). No obstant, la utilització de la tècnica de *top-down* MS obre la porta a poder fer un estudi més sofisticat de les protamines, incloent les diferents proteoformes i PTMs associades i superant les limitacions de les tècniques convencionals àcid – urea PAGE. D'aquesta manera, es podrà determinar millor el seu paper en la fertilitat i el desenvolupament embrionari.

Respecte a les PTMs de les protamines, poca cosa es coneix en comparació als rols proposats per les PTMs d'histones per a les que hi ha molta informació (Veure apartat 1.4.3.2.1). Tot i així, la PTM més estudiada en les protamines ha estat la fosforilació. Un cop sintetitzades, la P1 i les formes madures de P2 són àmpliament fosforilades en l'etapa d'espermàtida elongada per a poder unir-se correctament al DNA, per la *SRSF protein kinase 1* (SRPK1) i la *calcium/calmodulin dependent protein kinase IV* (CAMK4), respectivament (Oliva and Dixon, 1991; Green *et al.*, 1994; Pirhonen *et al.*, 1994c; Papoutsopoulou *et al.*, 1999; Wu *et al.*, 2000; Oliva, 2006; Carrell *et al.*, 2007; Jodar and Oliva, 2014). Tot i així, un cop les protamines es troben fortament unides al DNA, té lloc una desfosforilació extensa abans de la maduració espermàtica a l'epidídim (Gusse *et al.*, 1986; Carrell *et al.*, 2007), tot i que l'espermatozoide madur conserva alguns residus fosforilats (Veure annex 1) que són aportats en el moment de la fertilització (Gusse *et al.*, 1986; Pruslin *et al.*, 1987; Chirat *et al.*, 1993; Pirhonen *et al.*, 1994a; Papoutsopoulou *et al.*, 1999; Carrell *et al.*, 2007). Recentment, un estudi ha demostrat que els ratolins deficients en la xaperona *heat shock 70 kDa protein 4L* (Hspa4l), que és necessària en el reclutament de la fosfatasa *serine/threonine-protein phosphatase PP1-gamma catalytic subunit* (Ppp1cc2) a la cromatina espermàtica, són infèrtils, segurament degut a que Ppp1cc2 no desfosforila correctament P2 al testicle i provoca un increment anormal de la fosforilació de P2 a l'epidídim. En aquests ratolins, l'estat fèrtil és recuperat si el residu serina 56 (S56), que és un residu fosforilable, se substitueix per un residu alanina 56 (A56) que no es pot fosforilar (Itoh *et al.*, 2019). Un cop l'espermatozoide fertilitza l'oòcit, les protamines són ràpidament reemplaçades per histones

maternes, tot i que el mecanisme específic d'aquest recanvi ha estat poc estudiat (Jones *et al.*, 2012). En aquest sentit, s'ha demostrat de manera recent que just després de la fertilització, les protamines que empaqueten el DNA patern són extensament re-fosforilades per la SRPK1 d'origen matern i aquesta fosforilació marca l'inici de la reprogramació genòmica en què les protamines paternes seran reemplaçades per histones maternes (Gou *et al.*, 2020).

En la comparació d'individus normozoospèrmics que aconsegueixen embaràs amb individus que no l'aconsegueixen després de l'ús de FIV convencional, s'ha trobat en abundància alterada la quinasa encarregada de fosforilar protamina 1, la SRPK1 (Azpiazu *et al.*, 2014), suggerint el possible rol de les PTMs de protamines en les primeres etapes del desenvolupament. Amb les millores en les tècniques d'alt rendiment com la proteòmica, s'ha identificat múltiples combinacions de PTMs de les protamines en un *pool* d'espermatozoides de ratolí i en un *pool* d'espermatozoides humans (Brunner *et al.*, 2014; Castillo *et al.*, 2015) i s'ha proposat la hipòtesi de l'existència d'un codi de protamines com l'existent en histones que podria ser rellevant en les primeres etapes del desenvolupament embrionari (Brunner *et al.*, 2014). Concretament, en humans, s'han identificat mono-, di- i tri-fosforilacions, di-acetilacions i mono-metilació per P1 i per la família de P2, la forma madura HP3 amb una acetilació i una metilació, encara que no s'han pogut localitzar els residus modificats (Veure annex 1) (Castillo *et al.*, 2015). Tot i així, són necessàries millores en el protocol de purificació i en les estratègies proteòmiques, així com noves aproximacions avaluant individualment les PTMs de protamines per determinar la variabilitat interindividual en individus normozoospèrmics amb infertilitat d'origen desconegut i aclarir el seu paper en la fertilització i el desenvolupament embrionari primerenc.

La comparació del proteoma de l'espermatozoide de mostres de pacients normozoospèrmics amb infertilitat d'origen desconegut amb diferències en la taxa de fertilització, qualitat embrionària i/o la taxa d'èxit després de l'ús de TRAs pot ajudar a determinar les proteïnes de l'espermatozoide que poden tenir un rol en les primeres fases del desenvolupament embrionari (Taula 1.4) (Xu *et al.*, 2012; Zhu *et al.*, 2013; Azpiazu *et al.*, 2014; Frapsauce *et al.*, 2014; Légaré *et al.*, 2014; McReynolds *et al.*, 2014; Jodar *et al.*, 2018; Liu *et al.*, 2018; Wang *et al.*, 2018). Mitjançant proteòmica comparativa, es va concloure que alteracions en la capacitació i la reacció acrosòmica poden ser la causa de la fallida de la IIA en alguns casos (Xu *et al.*, 2012; Wang *et al.*, 2018). Tot i així, seria necessari també estudiar la possible desregulació de les PTMs, ja que s'ha demostrat que la fosforilació és essencial per la capacitació i la reacció acrosòmica (Castillo *et al.*, 2019). Per altra banda, la comparació d'individus normozoospèrmics que donen lloc a fallides de fertilització després de FIV convencional, ha contribuït a ampliar el coneixement de les possibles alteracions associades a la fertilització (Frapsauce *et al.*, 2014; Légaré *et al.*, 2014; Liu

et al., 2018). Concretament, s'han trobat en abundància alterada proteïnes relacionades amb la unió i fusió oòcit – espermatozoide com són la AKAP4, la *40S ribosomal protein SA* (RPSA) i la *zona pellucida-binding protein 1* (ZPBP1) (Frapsauce *et al.*, 2014; Légaré *et al.*, 2014; Liu *et al.*, 2018). No obstant, la comparació d'individus normozoospèrmics que pateixen fallides de fertilització després de FIV-ICSI ens pot donar més informació de les possibles proteïnes implicades en la inducció de l'activació oocitària per a què comenci el desenvolupament embrionari, atès que la tècnica FIV-ICSI se salta la majoria de les barreres biològiques implicades en la fertilització (Flaherty *et al.*, 1998). Recentment, un estudi col·laboratiu ha permès comparar per primer cop i mitjançant proteòmica amb marcatge diferencial, els espermatozoides de pacients que presenten elevada taxa de fertilització després de FIV-ICSI amb pacients que experimenten fallida de fertilització repetitiva després de FIV-ICSI (Veure annex 3). En aquest estudi, es van identificar 1400 proteïnes, de les quals 232 van ser considerades per a la seva anàlisi quantitativa, seguint els criteris estrictes descrits anteriorment (Barrachina *et al.*, 2019). L'anàlisi de grups va revelar que 5 proteïnes mitocondrials es trobaven sobreexpressades en el grup de fallida de fertilització. Així mateix, 4 proteïnes implicades, entre altres, en el funcionament del proteasoma es trobaven subexpressades en les mostres amb fallida de fertilització. Tots els resultats d'aquest estudi apunten que mecanismes com el metabolisme mitocondrial de l'espermatozoide podria tenir un paper rellevant, no només en la capacitat de l'espermatozoide per fertilitzar l'oòcit, sinó també en l'activació de l'oòcit (Veure annex 3).

La comparació del proteoma d'espermatozoides d'individus normozoospèrmics amb infertilitat d'origen desconegut segons la qualitat embrionària assolida després de l'ús de FIV-ICSI pot permetre valorar la participació de les proteïnes espermàtiques implicades en el desenvolupament embrionari primerenc, però no les implicades en el procés de fertilització. N'és un exemple, un estudi on es comparaven mostres d'espermatozoides d'individus normozoospèrmics amb la qualitat embrionària en l'estadi de blastocist i la taxa d'implantació que presentaven (Taula 1.4) (McReynolds *et al.*, 2014). Concretament, es van identificar 49 proteïnes diferencialment expressades que podrien implicar l'aportació d'un conjunt de proteïnes (*protein cargo*) presents en una abundància anòmala a l'embrió. Per exemple, la proteïna *testis-specific serine/threonine-protein kinase 2* (TSSK2) que es va trobar en abundància disminuïda en blastocists de baixa qualitat embrionària (McReynolds *et al.*, 2014). La proteïna TSSK2 es troba associada als centríols (Xu *et al.*, 2008) i ja que l'espermatozoide aporta el centríol proximal per a la formació del fus mitòtic per la primera divisió embrionària (Veure apartat 1.4.2), una alteració en proteïnes associades al centríol podria donar lloc a la baixa qualitat embrionària. Finalment, en l'estudi de Azpiazu et al., 2014, on es comparava la taxa d'embaràs després de FIV convencional, es va identificar en abundància disminuïda la proteïna *left-right determination factor 1* (LEFTY1) (Azpiazu *et al.*, 2014). Aquesta proteïna es troba

relacionada amb la determinació de la simetria esquerra-dreta dels òrgans durant el desenvolupament embrionari i també sembla ser important en la implantació embrionària (Raya and Izpisúa Belmonte, 2006; Tabibzadeh, 2011). En aquest mateix estudi, també es va identificar un augment d'histones en el subgrup que no aconseguia embaràs després de FIV convencional (Azpiazu *et al.*, 2014). De manera similar, un altre estudi va identificar un augment en diferents histones en el grup d'individus normozoospèrmics en què fracassava la FIV convencional (Légaré *et al.*, 2014; Liu *et al.*, 2018). Aquests resultats suggereixen una retenció incorrecta de les histones a l'espermatozoide madur que aporten un *cargo* de proteïnes amb una abundància alterada a l'embrió en el moment de la fertilització.

Taula 1.4 Estudis del contingut de proteïnes de l'espermatozoide humà madur. S'han incorporat estudis descriptius i comparatius segons els paràmetres seminals i la potencial contribució a l'embrió. S'han inclòs estudis de proteòmica.

Contingut de proteïnes	Estudi
Proteòmica descriptiva	
Compilació de les proteïnes de l'espermatozoide	(Amaral *et al.*, 2014a); (Wang *et al.*, 2016); (Jodar *et al.*, 2017); (Castillo *et al.*, 2018)
Espermatozoide sencer	(Johnston *et al.*, 2005); (Martínez-Heredia *et al.*, 2006); (Baker *et al.*, 2007); (de Mateo *et al.*, 2007); (Li *et al.*, 2007); (Wang *et al.*, 2013a)
Fraccionament subcel·lular de l'espermatozoide	(Kim *et al.*, 2007); (Naaby-Hansen and Herr, 2010); (Naaby-Hansen *et al.*, 2010); (de Mateo *et al.*, 2011a); (Gu *et al.*, 2011); (Nixon *et al.*, 2011); (Amaral *et al.*, 2013); (Baker *et al.*, 2013); (Kichine *et al.*, 2013); (Castillo *et al.*, 2014a)
PTMs	(Lefièvre *et al.*, 2007); (Vigodner *et al.*, 2013); (Wang *et al.*, 2013b); (Yu *et al.*, 2015); (Hetherington *et al.*, 2016); (Urizar-Arenaza *et al.*, 2019)
Detecció de *missing proteins*	(Jumeau *et al.*, 2015); (Vandenbrouck *et al.*, 2016); (Carapito *et al.*, 2017)
Proteòmica comparativa amb paràmetres seminals	
Normozoospèrmia amb infertilitat d'origen desconegut	(Selvam *et al.*, 2019)
Astenozoospèrmia	(Zhao *et al.*, 2007); (Martínez-Heredia *et al.*, 2008); (Chan *et al.*, 2009); (Siva *et al.*, 2010); (Chao *et al.*, 2011); (Parte *et al.*, 2012); (Shen *et al.*, 2013); (Amaral *et al.*, 2014b); (Hashemitabar *et al.*, 2015); (Liu *et al.*, 2015); (Saraswat *et al.*, 2017); (Nowicka-Bauer *et al.*, 2018); (Guo *et al.*, 2019); (Moscatelli *et al.*, 2019); (Sinha *et al.*, 2019)
Oligozoospèrmia	(Botta *et al.*, 2009)
Teratozoospèrmia	(Liao *et al.*, 2009)
Oligoastenoteratozoospèrmia	(Thacker *et al.*, 2011)
Proteòmica comparativa per determinar la contribució de l'espermatozoide a l'embrió	
Capacitació i reacció acrosòmica amb o sense PTMs	(Ficarro *et al.*, 2003); (Secciani *et al.*, 2009); (Wang *et al.*, 2015); (Sun *et al.*, 2014); (Redgrove *et al.*, 2011); (Castillo *et al.*, 2019); (Hernández-Silva *et al.*, 2020)
Fertilització	(Pixton *et al.*, 2004); (Frapsauce *et al.*, 2009); (Frapsauce *et al.*, 2014); (Légaré *et al.*, 2014); (Liu *et al.*, 2018)
Qualitat embrionària i èxit de les TRAs	(Xu *et al.*, 2012); (Zhu *et al.*, 2013); (Azpiazu *et al.*, 2014); (McReynolds *et al.*, 2014); (Wang *et al.*, 2018)

PTMs: Modificacions posttraduccionals; TRAs: Tècniques de Reproducció Assistida

2.

HIPÒTESI I OBJECTIUS

2. HIPÒTESI I OBJECTIUS

2.1. Hipòtesi

L'aplicació de tècniques d'alt rendiment, com la transcriptòmica i la proteòmica, en l'estudi de l'espermatozoide humà, ha permès obtenir grans avenços en la coneixença de la seva composició i funció normal, així com d'algunes alteracions associades a la infertilitat. A més, s'ha demostrat en els últims anys, que l'espermatozoide, a part d'aportar a l'embrió el DNA amb marques epigenètiques, també aporta una complexa població de proteïnes i RNAs, algunes d'elles crucials per la fertilització i el l'embriogènesi. Tot i així, encara manquen estudis per determinar l'impacte de la participació paterna a l'embrió en desenvolupament. L'estudi de pacients normozoospèrmics amb infertilitat d'origen desconegut és clau, ja que no presenten defectes en l'espermatogènesi i permet determinar, caracteritzar i analitzar la correcta contribució de la cèl·lula espermàtica al zigot i identificar potencials alteracions. En aquesta mateixa línia, es poden distingir dos tipus d'aproximacions quan s'utilitzen tècniques d'alt rendiment: els estudis descriptius amb inferència *in silico* i els estudis comparatius.

S'ha proposat que el contingut d'RNAs de l'espermatozoide contribueixen al desenvolupament embrionari primerenc, sobretot els RNAs no codificants que regulen l'expressió gènica. Recentment, s'ha descrit un nou tipus d'RNA no codificant en altres tipus cel·lulars, els RNAs circulars (circRNAs), que regulen l'expressió gènica tan a nivell transcripcional com posttranscripcional. L'espermatozoide és una cèl·lula molt especial i diferenciada: no transcriu ni tradueix, presenta una estructura de la cromatina molt particular i posseeix una població abundant, complexa i estable d'RNAs tan codificants com no codificants. El fet que els circRNAs presenten vides mitjanes llargues i poden regular l'expressió gènica, suggereix que l'espermatozoide podria tenir una població única de circRNAs. De fet, degut a que els circRNAs no presenten extrem 3', es podrien estar protegint a ells mateixos o a altres RNAs no codificants de la degradació durant el trànsit pel tracte sexual masculí i femení i, per tant, podrien ser claus per la contribució paterna al desenvolupament embrionari primerenc.

Per altra banda, també s'ha descrit el rol en el desenvolupament embrionari de les PTMs de les histones que empaqueten d'un 5 – 15 % del DNA patern. Tot i així, les protamines, que empaqueten de forma majoritària el genoma de l'espermatozoide (85 – 95 %), i les seves potencials PTMs han estat poc explorades. Per tant, podria ser que existís un codi de PTMs de les protamines com el que existeix amb les histones que podria contribuir a l'embriogènesi.

En els pacients normozoospèrmics amb infertilitat d'origen desconegut, la predicció de la taxa d'èxit de les TRAs i la contribució de l'espermatozoide a la qualitat embrionària ha estat poc explorat. A nivell proteic, s'ha determinat la contribució de l'espermatozoide en embrions que provenen de IIA o FIV convencional i s'han relacionat defectes espermàtics associats amb la capacitació i la fertilització que podrien ser la causa latent de la infertilitat en aquests casos. Per contra, pocs estudis han relacionat la qualitat embrionària després de FIV-ICSI en individus normozoospèrmics amb infertilitat d'origen desconegut. Es creu que l'espermatozoide té un paper fonamental durant els primers estadis del desenvolupament embrionari primerenc, abans de l'activació del genoma que té lloc entre els dies 2 i 3 entre l'estadi de 4 – 8 cèl·lules. Per aquest motiu, manquen estudis que valorin la qualitat embrionària en dia 2 d'embrions provinents d'individus normozoospèrmics amb infertilitat d'origen desconegut després de FIV-ICSI per poder determinar la possible contribució paterna a l'embriogènesi i poder predir la taxa d'èxit de les TRA.

Per tant, la hipòtesi d'aquesta Tesi Doctoral és que la combinació tan d'estudis descriptius amb inferència *in silico* com comparatius a diferents nivells moleculars de l'espermatozoide humà permetrà augmentar la comprensió de la contribució paterna a l'embrió. Per aquest motiu, s'espera que en individus normozoospèrmics amb infertilitat d'origen desconegut, la descripció de la població de circRNAs i de les PTMs presents a les protamines de l'espermatozoide humà, així com la comparació d'alteracions en el conjunt de proteïnes de l'espermatozoide que afectin a la qualitat embrionària, permetrà aclarir la normalitat i les alteracions de la contribució paterna en la fertilització i les primeres etapes del desenvolupament embrionari.

2.2. Objectius

Per tant, els objectius específics i detallats d'aquesta Tesi Doctoral que es deriven de la hipòtesi establerta (Veure apartat 2.1) són els següents:

1. Determinar el perfil normal d'RNAs circulars (circRNAs) de l'espermatozoide madur en pacients normozoospèrmics que visiten la clínica de reproducció per una avaluació de la fertilitat en ambdós membres de la parella, mitjançant RNA-seq.

2. Determinar el perfil normal de les modificacions posttraduccionals (PTMs) presents en les protamines de l'espermatozoide madur, mitjançant cromatografia líquida seguit d'espectrometria de masses. Concretament:

    2.1. Millorar el protocol d'aïllament de protamines per a experiments d'espectrometria de masses.

    2.2. Determinar el perfil normal de les PTMs presents en les protamines de l'espermatozoide en individus normozoospèrmics que visiten la clínica de reproducció per una avaluació de la fertilitat en ambdós membres de la parella.

3. Identificar les alteracions en el conjunt de proteïnes de l'espermatozoide en pacients normozoospèrmics amb infertilitat d'origen desconegut amb elevada o baixa qualitat embrionària en el dia 2 del desenvolupament embrionari després de fecundació *in vitro* amb injecció intracitoplasmàtica (FIV-ICSI), mitjançant espectrometria de masses.

3.

TREBALLS

3.1. Treball 1

RNAs circulars de l'espermatozoide: Tipus, patrons de resistència a la RNasa R i potencial epigenètic

En preparació

Sperm circular RNAs: Types, RNase R resistance patterns and epigenetic potential.
Soler-Ventura A, Odriozola A, Castillo J, Delgado-Dueñas D, Guimerà M, Corral JM, Ballescà JLL, Oliva R, and Jodar M.

RNAs circulars de l'espermatozoide: Tipus, patrons de resistència a la RNasa R i potencial epigenètic

Objectiu: Caracteritzar els RNAs circulars (circRNAs) de l'espermatozoide humà madur i determinar si els elements intrònics (IEs) són, de fet, circRNAs intrònics. A més, posar a punt el tractament amb RNasa R amb RNA d'espermatozoide i així validar per primer cop alguns dels circRNAs predits tan en aquesta cèl·lula com en el testicle. Finalment, inferir també les funcions dels circRNAs de l'espermatozoide relacionades amb la fertilitat i l'embriogènesi.

Metodologia: Es va realitzar RNA-seq de mostres d'espermatozoides d'individus normozoospèrmics. L'algoritme KNIFE va permetre predir els circRNAs exònics de l'espermatozoide. Es va posar a punt i es van testar diferents candidats al tractament amb RNasa R i, es va determinar si la descircularització d'algun d'ells era específica o aleatòria. També es van inferir potencials funcions *in silico*.

Resultats: L'espermatozoide conté una població complexa de circRNAs exònics i, els IEs descrits prèviament són circRNAs intrònics. Els circRNAs predits del gàmeta masculí van ser classificats en tres grups segons la seva resistència al tractament amb RNasa R. Al contrari que els circRNAs presents a les cèl·lules somàtiques, els circRNAs de l'espermatozoide sembla que es troben majoritàriament en la seva forma circular, sense la presència dels seus equivalents lineals. L'estabilitat d'aquests circRNAs suggereix que la seva resistència al tractament amb RNasa R podria ser utilitzat com un test d'integritat de l'RNA d'espermatozoide, similar a l'ús del *RNA integrity number* (RIN) en cèl·lules somàtiques. També es van identificar alguns circRNAs sensibles al tractament amb RNasa R degut a una descircularització no aleatòria. A nivell funcional, els circRNA espermàtics sembla que emergeixen principalment de gens enriquits en funcions epigenètiques. Específicament, alguns circRNAs de l'espermatozoide es van associar a *sperm RNA elements* (SREs) necessaris per aconseguir un naixement i/o es van trobar alterats en pacients obesos. Similarment, alguns circRNAs embrionaris semblen ser aportats específicament per via paterna. Finalment, la presència de *open reading frames* (ORFs) va revelar un potencial traduccional per algun d'ells.

Conclusions: L'espermatozoide madur conté una població complexa de circRNAs, incloent circRNAs exònics i intrònics. Els resultats obtinguts en testicle donen suport a que durant l'espermatogènesi, les formes lineals són gradualment degradades, mentre que els circRNAs són retinguts selectivament a l'espermatozoide madur. Alguns circRNA del zigot són aportats per via paterna, així com alguns SREs semblen trobar-se associats a circRNAs, donant suport a que els circRNAs podrien ser necessaris pel correcte funcionament de l'espermatozoide i tenir un paper més enllà de la fertilització. La presència d'ORFs intactes en els circRNAs de l'espermatozoide reforça la hipòtesi que els circRNAs del gàmeta masculí podrien ser traduïts a pèptids i presentar rols putatius a l'embriogènesi.

Article in preparation

Sperm circular RNAs: Types, RNase R resistance patterns and epigenetic potential.

Ada Soler-Ventura[1], Ainitze Odriozola[1], Judit Castillo[1], David Delgado-Dueñas[1], Marta Guimerà[2], Juan Manuel Corral[3], Josep Lluís Ballescà[1,2], Rafael Oliva[1,4*] and Meritxell Jodar[1,4*]

[1]Molecular Biology of Reproduction and Development Research Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, 08036, Spain. [2]Clinic Institute of Gynecology, Obstetrics and Neonatology (ICGON), Hospital Clínic, Barcelona, 08036, Spain. [3]Department of Urology, Hospital Clínic, Barcelona, 08036, Spain. [4]Biochemistry and Molecular Genetics Service, Hospital Clínic, Barcelona, 08036, Spain. * To whom correspondence should be addressed. Tel: +34 93 402 18 77; Fax: +34 93 403 52 78; Email: meritxell.jodar@ub.edu. Correspondence may also be addressed to roliva@ub.edu

ABSTRACT

Circular RNAs (circRNAs) are a widely distributed new type of non-coding RNAs with important regulatory roles. CircRNAs are synthesized by backsplicing and lack the 3'-free end, which confer higher stability and longer half-lives compared to linear RNAs. Here, we have characterized and experimentally validated the circRNAs content in the human male gamete. Our results show that the spermatozoon contains a complex and stable population of exonic and intronic circRNAs, classified in three groups according to their RNase R resistance. In contrast to the circRNAs from somatic cells, sperm circRNAs have been mainly found in their circular form, with no presence of their linear cognates. Surprisingly, we also identified a proportion of circRNAs sensible to the RNase R treatment due to a non-stochastic decircularization. At the functional level, sperm circRNAs seem to arise mainly from genes enriched in epigenetic functions. In particular, some sperm circRNAs were associated to sperm RNA elements required to achieve a live birth and/or found altered in obese patients. Likewise, some early embryonic circRNAs appear to be specifically provided by the sperm and some of them are potentially translated to functional peptides. Altogether, these results suggest crucial roles for sperm circRNAs in fertility, early embryogenesis and beyond.

INTRODUCTION

In the last years, circular RNAs (circRNAs) have emerged as a new type of non-coding RNA (ncRNA) with important associations in multiple types of diseases, including cancer (Salzman et al., 2012; Jeck et al., 2013; Memczak et al., 2013). The circularization of these RNAs is produced by backsplicing, a covalent bond between a 5' splice donor and a 3' splice acceptor (Barrett and Salzman, 2016; Patop et al., 2019). It has been postulated that the RNA polymerase II transcribes circRNAs and the canonical splicing machinery is probably involved in their production. Moreover, exon-skipping events or flanking longer-than-average introns containing inverted repeats – in human mostly ALU repeats – seem to facilitate circularization by complementarity (Jeck et al., 2013; Liang and Wilusz, 2014; Ivanov et al., 2015). Likewise, several studies have demonstrated the involvement of RNA-binding proteins (RBPs) in the regulation of circRNA production, which includes the splicing factor muscleblind (MBL), the protein quaking (QKI), the RNA-binding protein FUS (FUS), and serine/arginine-rich proteins. In contrast, the double-stranded RNA-specific adenosine deaminase (ADAR) and the DExH-Box helicase 9 (DHX9) have been suggested as mediator of the suppression of circRNAs production by blocking the looping of intronic complementary sequences (Ashwal-Fluss et al., 2014; Conn et al., 2015; Kramer et al., 2015; Rybak-Wolf et al., 2015; Aktaş et al., 2017; Errichelli et al., 2017; Li et al., 2017; Yu et al., 2017).

The resulting circRNAs lack the 3'-end, which confer higher stability than linear forms, due to their resistance to endo- and exonucleases-mediated degradation, and longer half-lives than their linear RNA cognates (Jeck et al., 2013; Enuka et al., 2016; Szabo and Salzman, 2016). CircRNAs can originate from exons (exonic circRNAs), from introns (intronic circRNAs), from exons and introns (exon-intron circRNAs) or from intergenic regions and antisense locations (Memczak et al., 2013; Barrett and Salzman, 2016). However, circRNAs are thought to be mostly derived from exons of protein-coding genes and tend to exclude the first and the last exon of the corresponding gene (Salzman et al., 2012; Jeck et al., 2013; Aufiero et al., 2018). In addition, exonic circRNAs appear to be exported to the cytoplasm through the RNA helicases UAP56 and URH49 in a size-dependent manner (Huang et al., 2018), while exon – intron and intronic circRNAs seem to remain inside the nucleus (Jeck et al., 2013; Zhang et al., 2013; Li et al., 2015; Huang et al., 2018).

It has been observed that circRNAs are differentially expressed depending on tissue and subcellular localizations, as well as on the developmental stage (Jeck et al., 2013; Salzman et al., 2013; Westholm et al., 2014; Rybak-Wolf et al., 2015). Diverse circRNA populations have been identified in several species, such as humans, for which circRNAs have been described in multiple tissues and cell types, including testis (Westholm et al.,

2014; Rybak-Wolf et al., 2015; Venø et al., 2015; Dong et al., 2016; Gruner et al., 2016; Sun et al., 2016). Additionally, circRNAs are suggested to be key regulators for organogenesis and cell differentiation during embryo development (Dang et al., 2016; Huang et al., 2019) and have been postulated to be connected to the development of different pathologies, such as cancer and cardiovascular diseases, due to their role in gene expression regulation (reviewed elsewhere (Haddad and Lorenzen, 2019)).

CircRNAs could have important regulatory roles at both transcriptional and post-transcriptional levels. To date, the most popular function associated with these circRNAs correspond to their ability for microRNA (miRNA) sponging. It has been shown that the circRNA ciRS-7 contains more than 70 binding sites for the mirR-7 (Hansen et al., 2013; Memczak et al., 2013). However, recent findings concluded that most of the other circRNAs do not contain more miRNA binding sites than randomly expected. Therefore, the high number of miRNA binding sites found in ciRS-7 could be less frequent in circRNAs than previously thought (Guo et al., 2014; Dori and Bicciato, 2019; Kristensen et al., 2019). Other described functions for circRNAs are protein sponging of RBPs, protein enhancing, protein scaffolding and protein recruitment to a subcellular localization (Kramer et al., 2015; Kristensen et al., 2019). Additionally, circRNAs can regulate alternative splicing (Zhang et al., 2014; Starke et al., 2015) and modulate the RNA polymerase II activity (Zhang et al., 2013). Recently, some studies have also postulated that specific circRNAs containing internal ribosome entry sites (IREs), large intact ORFs and N6-methyladenosine (m6A)-modified start codons could be translated in a cap-independent manner to peptides, and could potentially act as functional proteins (Wang and Wang, 2015; Pamudurti et al., 2017; Yang et al., 2017; Kristensen et al., 2019; Tang et al., 2020).

While several bioinformatics tools have been developed for the in-silico prediction of exonic circRNAs from RNA-seq data, prediction tools for intronic circRNAs are not as well established (Szabo and Salzman, 2016; Zeng et al., 2017). In any case, any prediction of circRNAs should always be accompanied by a robust experimental validation, since RNA-seq protocols introduce technical artifacts that can result in the identification of false-positive matches (Barrett and Salzman, 2016). Reverse transcription and the subsequent PCR can also produce false-positive detections due to RT-PCR artifacts scrambled junctions' misidentification, such as template switching, PCR bias attributable to different amplification efficiencies and rolling circle amplification. Likewise, mRNAs from circular DNAs, exon repeats and trans-splicing events could derive on a false circRNA prediction (Szabo and Salzman, 2016; Iparraguirre et al., 2019; Patop et al., 2019). The robust validation of the circRNA in-silico prediction must first include the identification of the circularization point, by the specific amplification using

divergent primers and the direct sequencing of the circRNA junction. After that, it is necessary to test the resistance of both the circRNA and its cognate linear isoform to the RNase R, which is an exoribonuclease that hydrolyzes the free 3'-end of the RNAs, degrading the linear RNAs and leaving the circRNAs intact (Salzman et al., 2012; Szabo et al., 2015; Kristensen et al., 2019; Patop et al., 2019).

In the specific case of the male gamete, high-throughput techniques have allowed characterizing a complex population of RNAs, which are specifically retained in this cell (Ostermeier et al., 2004; Jodar et al., 2013). In particular, spermatozoa contain coding and non-coding RNAs that include fragmented and intact mRNAs, fragmented ribosomal (r) RNAs, miRNAs, tRNA-derived sncRNAs (tRFs), small interfering (si) RNAs, long non-coding (lnc) RNAs, YRNAs, small nuclear (sn) RNAs, small nucleolar (sno) RNAs, and P element-induced wimpy testes (PIWI)-interacting (pi) RNAs (Krawetz et al., 2011; Jodar et al., 2013; Gòdia et al., 2018; Jodar, 2019). It is interesting to highlight that while some of the sperm RNAs have also been identified in other cell types, others seem to be sperm-specific. For instance, certain RNAs derived from introns and intergenic regions, named as intronic or intergenic elements (IEs and Blips, respectively), are abundantly and specifically contained in the mature spermatozoa, although their function has not been elucidated yet (Jodar et al., 2013, 2015).

Of note, the abundance of circRNAs appears to show a strong negative correlation with the replication rate of the cells. This would suggest that non-replicative cells, such as the sperm, tend to accumulate a larger population of circRNAs (Bachmayr-Heyda et al., 2015). In fact, a switch from linear to circular forms taking place during the late steps of murine spermatogenesis has recently been described (Tang et al., 2020). These findings, together with the facts that the male gamete harbors a unique chromatin structure, is transcriptionally and translationally inert and possesses a complex population of RNAs that includes IEs, lead us to hypothesize that circRNAs, which show higher resistance and stability than their linear cognates, would be abundant in the RNA profile of the mature sperm. Indeed, some recent studies have started to characterize this novel RNA population in male germ cells and mature spermatozoa (Chioccarelli et al., 2019; Ragusa et al., 2019; Gòdia et al., 2020; Tang et al., 2020). However, the study of sperm circRNAs is still in its infancy, and additional studies are needed in order to robustly characterize them and unravel their potential functionality. Therefore, in this study, we have conducted an in-depth characterization of the circRNAs contained in the mature human sperm. As a novelty, we have included a robust validation of sperm circRNAs and determined whether the abundant sperm IEs are, in fact, intronic circRNAs. Additionally, we have set-up and tested the resistance of some sperm

circRNAs to the RNase R treatment, and inferred their potential functions related to fertility and early embryogenesis.

MATERIAL AND METHODS

The overall methodology employed in this study is depicted in Supplementary Figure 1.

Reagents

PureSperm® (Catalog # PSB-100, # PS100-100) was purchased from NidaCon International AB (Gothenburg, Sweden). CryoSperm™ (Catalog # 11010010) was purchased from Origio (Målov, Denmark). RNeasy Mini Kit protocol (Catalog # 74104) and MiniElute PCR Purification kit (Catalog # 28004) were purchased from Qiagen (Hilden, Germany). RNA*later*® (Catalog # R0901-100ML) and SeqPlex RNA Amplification kit (Catalog # SEQR-10RXN) were purchased from Sigma-Aldrich (San Luis, Missouri, US). PowerUp™ SYBR™ Green Master Mix (Catalog # A25741) was purchased from Applied Biosystems (Foster City, California, US). Turbo DNase free kit (Catalog # AM1907), Quant-iT™ RiboGreen™ RNA Assay Kit (Catalog # R11490), Quant-iT™ dsDNA Assay Kit (Catalog # P7581), SuperScript™ III Reverse Transcriptase (Catalog # 18080093), oligo (dT) primers (Catalog # AM5730G), Random Hexamers (Catalog # N8080127) and Maxima Reverse Transcriptase (Catalog # EP0741) were purchased from Thermo Fisher Scientific (Waltham, Massachusetts, US). NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (Catalog # E7370S) and NEBNext Multiplex Oligos for Illumina (Index Primers Set 1) (Catalog # E7335S) were purchased from New England Biolabs (Ipswich, Massachusetts, US). AmpureXP (Catalog # A63880) was purchased from Beckman Coulter (Brea, California, US). KAPA Library Quantification Kit (Catalog # 7960204001) was purchased from Roche (Basel, Switzerland). RNase R (Catalog # RNR07250) was purchased from Epicentre Technologies Corp (Madison, Wisconsin, US).

Biological resources

Human semen samples (n=12; 3 for RNA-seq analysis and 9 for experimental validation) were obtained from patients undergoing routine semen analysis at the Assisted Reproduction Unit from the Clinic Institute of Gynecology, Obstetrics, and Neonatology, from the Hospital Clínic de Barcelona (Barcelona, Spain). The human testis biopsy was collected during a routine testicular sperm extraction. All samples were used in accordance with the appropriate ethical guidelines and Internal Review Board, and the biological material storing and processing was approved by the Clinical Research Ethics Committee of the Hospital Clínic de Barcelona (Barcelona, Spain). All participants have

provided written informed consent under the Declaration of Helsinki. Human embryonic kidney (HEK) cells were collected from a confluent cellular culture.

Sample collection and preparation

The ejaculates were collected by masturbation into sterile containers after 3-5 days of sexual abstinence. The evaluation of the seminal parameters was performed using the automatic semen analysis system CASA (Proiser, Paterna, Spain) and patients were classified as normozoospermic according to the World Health Organization (WHO) guidelines (World Health Organization, 2010). In order to purify spermatozoa from other cell types, a 50 % density PureSperm® gradient separation (NidaCon International AB) was performed according to manufacturer's recommendations. All samples were verified to contain < 1 % of somatic contamination. Subsequently, samples were cryopreserved with CryoSpermTM (Origio), according to the manufacturer's instructions, and kept in liquid nitrogen. Cryopreserved sperm samples were thawed, centrifuged, resuspended in RLT buffer (Qiagen) supplemented with 1.5 % β-mercaptoethanol and stored at -80ºC until further processing. The testis biopsy was kept in RNAlater® (Sigma-Aldrich) until further processing. For the HEK cells, the culture medium was removed by centrifugation and cells were stored at -80 ºC in RLT buffer (Qiagen) supplemented with 1.5 % β-mercaptoethanol until further processing.

RNA isolation

RNA from human testis was extracted with the RNeasy Mini Kit protocol (Qiagen) following the manufacturer's instructions. The long RNA fraction from sperm and HEK cells were individually extracted following the RNeasy Mini Kit protocol (Qiagen), with some modifications previously established in Goodrich et al., 2013 (Goodrich et al., 2013). All RNA samples were treated with the Turbo DNase free kit (Thermo Fisher Scientific) to remove potential DNA contamination. The absence of any residual DNA was assessed by PowerUp™ SYBR™ Green Master Mix real-time PCR (Applied Biosystems), using protamine 1 (*PRM1*) specific primers (Supplementary Table 1). Extracted RNA was quantified with Quant-iT™ RiboGreen™ RNA Assay Kit (Thermo Fisher Scientific). For sperm samples, the RNA integrity and the absence of somatic RNA contamination was evaluated by conducting reverse transcription of 10 ng of sperm RNA using SuperScript™ III Reverse Transcriptase (Thermo Fisher Scientific) and oligo (dT) primers (Thermo Fisher Scientific). Subsequently, a real-time PCR using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems) was performed targeting *PRM1* and the leukocyte-specific marker Protein Tyrosine Phosphatase Receptor Type C (*PTPRC*) (Supplementary Table 1). For testis and HEK samples, the integrity of the

RNA was corroborated by assessing the RNA integrity number (RIN) through the TapeStation 4200 (Agilent Technologies, Santa Clara, California, US), which resulted in high-quality RIN values of 6.6 (White et al., 2018) for the testis sample and > 9.5 for HEK cells.

RNA sequencing

Once confirmed the good quality of the RNA extracted from the sperm samples, 5 ng of each biological replicate (n=3) were reverse transcribed and the cDNA was amplified with SeqPlex RNA Amplification (Sigma-Aldrich). Residual nucleotides and primer dimers were removed with the MiniElute PCR Purification kit (Qiagen), following the manufacturer's instructions. Library construction was performed with 50 ng of amplified cDNA, previously quantified using Quant-iT™ dsDNA Assay Kit (Thermo Fisher Scientific), and the NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (New England Biolabs). NEBNext Multiplex Oligos for Illumina (Index Primers Set 1) (New England Biolabs) were included for sample barcoding. The surplus of primer dimers was cleaned-up with AmpureXP (Beckman Coulter). The correct elimination of primer dimers was assessed with the DNA High Sensitivity Assay in the Bioanalyzer 2100 (Agilent Technologies). Libraries were quantified through KAPA Library Quantification Kit (Roche) and pooled. Paired-end sequencing was performed for 2x75 base pairs (bp) using the Illumina HiSeq-2500 sequencer.

Exonic circRNAs prediction

In order to predict exonic circRNAs in human sperm, oocyte and zygote, the KNIFE algorithm created by Szabo and colleagues (Szabo et al., 2015) was applied using the cancer genomics cloud platform (http://www.cancergenomicscloud.org/) (Lau et al., 2017). Specifically, the in-house raw RNA-seq data from the spermatozoa samples included in this study (n=3), as well as public raw RNA-seq datasets from Dang et al., 2016 (Dang et al., 2016), which includes single-cell oocytes (n=4) and single-cell zygotes (n=4), were used. To apply the KNIFE algorithm, fastq files were mapped using the Bowtie2 to indexed reference genome (h38), ribosomal RNA, and databases of linear exon-exon junctions and scrambled exon-exon junctions. Briefly, the KNIFE algorithm calculates a statistical confidence score for each read based on the alignment properties, including mismatches and mapping quality to reduce the number of false-positives. Furthermore, it applies a two-step logistic generalized linear model (GLM) to determine whether a junction matching to a circRNA is a true positive or an artefactual alignment, giving a high-confident p value score (Szabo et al., 2015). KNIFE combined reports were employed to identify and quantify exonic circRNAs splicing events and strict criteria were

applied to consider true positive circRNAs, namely ≥ 2 reads containing a circularized junction with a p value score ≥ 0.9 (Szabo et al., 2015). Of note, the comparison of the KNIFE algorithm with other circRNAs prediction tools concluded that KNIFE algorithm reached the best-balanced prediction between precision and sensitivity (Zeng et al., 2017).

Selection of sperm *intronic elements* for circRNA validation

The most abundant IEs contained in the spermatozoa were obtained from the list provided in Jodar et al., 2015 (Jodar et al., 2015). Data were extracted and updated to the hg38 genome (Supplementary Table 2). Only those IEs over the 0.995 percentile rank and expressing only one intron were selected. All the candidates that after the hg19 to hg38 conversion aligned to not expected regions, showed reads in a very small region of the intron or in the adjacent exons, or had intronic regions containing antisense RNAs and lncRNAs were excluded for further analysis.

Validation of the circRNAs candidates

In order to validate the circRNAs candidates, the circular specific junction was first determined. To that end, convergent and divergent primers were designed using the Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/) (Koressaar and Remm, 2007; Untergasser et al., 2012) (Supplementary Table 1). The amplification products were visualized on a 3 % agarose gel and analyzed by direct sequencing. Subsequently, the circularity of the selected RNAs was validated by assessing the resistance to the RNase R treatment. The RNase R resistance test (Epicentre Technologies Corp.) was first optimized for its application to sperm RNA by treating 100 ng of RNA from either sperm or HEK cells with decreasing Units (U) of RNase R (5 U, 1 U, 0.5 U or 0.1 U), during 30 min at 37 ºC. As negative controls, cells subjected to the same treatment but without RNase R were also included (mock treatment). The cDNA was synthesized using Random Hexamers (Thermo Fisher Scientific) and Maxima Reverse Transcriptase (Thermo Fisher Scientific). To determine the efficacy of the RNase R treatment, a qPCR amplification was carried out using convergent and divergent primers targeting RNAs that are specific to each cell type. In particular, the Calmodulin Regulated Spectrin Associated Protein 1 (*CAMSAP1*) transcript, previously described in the literature (Salzman et al., 2013), was selected for the evaluation of HEK cells. To the best of our knowledge the resistance to the RNase R treatment for sperm RNA was never assessed before, therefore we selected the abundant sperm specific *PRM1* transcript for its evaluation. However, no *PRM1* circular forms were detected using divergent primers (data not shown). For this reason, *PRM1* amplification with convergent primers were

used to test the efficiency of the RNase R treatment for sperm RNAs, but we also included the amplification of *TRIM66* with divergent primers, which is one of the possible selected sperm intronic circRNAs predicted herein, to test the resistance of sperm circRNA to RNase R treatment. The resistance to the RNase R treatment for each specific product was determined by the difference between the cycle threshold (ΔCt) of the mock treatment (the no treated sample) and the RNase R treatment (ΔCt = Ct mock treatment – Ct RNase R treatment). The values were graphically represented by the $\log_2^{(2\Delta Ct)}$. Once the best conditions were established, the resistance to the RNase R treatment of 3 intronic and 5 exonic circRNAs was tested in sperm (n=9) and testis (n=1) samples. In addition, the amplification of *TRIM66* with divergent primers and *PRM1* with convergent primers were also assessed in all the samples as controls for RNA degradation and RNase R treatment efficiency, respectively. The primer sets used in this assay are listed in Supplementary Table 1.

Validation of specific circRNAs decircularization

The possible presence of a specific point of decircularization was suggested after detecting that certain confirmed sperm circRNAs were sensible to the RNase R treatment. *VRK1* was selected as a representative of sperm circRNAs showing degradation after RNase R treatment. To check whether this decircularization was the result of either a random degradation process or a specific cleavage, several sets of divergent primers covering the 10 exons (from exon 2 to exon 11) of the circRNA derived from *VRK1* gene were designed (Supplementary Table 3; Supplementary Figure 2A). The amplification efficiency and the circular specific junction were verified by direct sequencing of the amplified products. Subsequently, the RNase R treatment was carried out in 5 different biological replicates to test the resistance of the products obtained. The difference between the resistance to the RNase R treatment of the amplified products and the negative control was analyzed to delimitate the decircularization region. To verify the efficiency of the methodology, the same procedure was performed with 4 biological replicates for the circRNA derived from the exon 2 of *ZNF608* gene (Supplementary Table 3; Supplementary Figure 2B), which showed resistance to the RNase R treatment. The values were graphically represented by the $\log_2(2^{\Delta Ct})$.

Functional and epigenetic potential prediction

The gene names corresponding to the list of exonic circRNAs commonly found in the three sperm biological replicates were uploaded to the Gene Ontology Consortium database (http://www.geneontology.org/) (Ashburner et al., 2000; The Gene Ontology Consortium, 2019), based on PANTHER v13.1 database (Release date 2019 – 12 – 09),

to predict the potential biological functions of the sperm circRNAs. The significance of enrichment analyses was calculated by a Fisher's exact test. P-values < 0.05 after Bonferroni adjustment were considered statistically significant. Potential epigenetic functions for the predicted exonic circRNAs were assessed through the dataset available at http://epifactors.autosome.ru/ (Medvedeva et al., 2015).

The list provided in Jodar et al., 2015 (Jodar et al., 2015) of the 648 sperm RNA elements (SREs) required to achieve a live birth were employed to infer the sperm circRNAs potential role in male fertility. Similarly, the 487 SREs associated with obesity and/or with an epigenetic potential listed in Swanson et al., 2020 (Swanson et al., 2020) were used to extrapolate potential sperm circRNAs linked to body mass index (BMI), which could alter sperm functionality. Only those SREs embedded inside a predicted exonic circRNA were considered.

After applying KNIFE algorithm (Szabo et al., 2015) to the data from Dang et al., 2016 corresponding to human oocytes and zygotes (Dang et al., 2016), we searched for those exonic circRNAs found in sperm and zygote and absent in the oocyte to infer a potential paternal origin of embryonic circRNAs, while discarding a maternal origin. To this end, we searched those exonic circRNAs detected in all three sperm biological replicates included in this study that were contained in all the zygote biological and technical replicates and not detected in any of the biological nor the technical oocyte replicates. Only exonic predicted circRNAs with ≥ 2 reads and p value score ≥ 0.9 were considered for the analysis. The overlapping between sperm and zygote exonic circRNAs were named here paternally-derived embryonic circRNAs.

Computational prediction of 7-mer repetitions and translational potential

The prediction of 7-mer miRNA seeds was conducted on the validated exonic and IE circRNAs, as well as in the paternally-derived embryonic circRNAs, by applying an in-house R script (https://www.r-project.org/) (Development Core Team, 2018), in which 7 nt-repetitions, corresponding to potential miRNAs binding sites, were counted in the full circRNA sequence.

The prediction of the translational potential on the validated circRNAs and the paternally-derived embryonic circRNAs was performed using the Open Reading Frame Finder (https://www.ncbi.nlm.nih.gov/orffinder/) (Wheeler et al., 2003), which searches for open reading frames (ORFs) in the DNA sequence. The following parameters were applied: 75 nt of minimal ORF length, standard genetic code, and "ATG" for ORF start codon to use. "Clustal O" alignment (http://www.ebi.ac.uk/Tools/msa/clustalo/) (Goujon et al., 2010; Sievers et al., 2011) was conducted to compare the potential peptide sequences against the canonical ones.

RESULTS

Spermatozoa contain a complex population of circRNAs

KNIFE algorithm predicted a total of 5,830 unique exonic circRNAs from 2,824 different genes in sperm, distributed as follows: 3,242 circRNAs from 1,871 different genes in sample 1, 3,449 circRNAs from 1,948 different genes in sample 2, and 1,898 circRNAs from 1,301 different genes in sample 3 (Figure 1A). From those, 737 exonic circRNAs derived from 545 different genes overlapped between biological replicates and were considered for further analyses (Supplementary Table 4; Figure 1A). The most abundantly identified exonic circRNA in all sperm samples arised from *LINC02050,* with an average of almost 5,600 reads. Looking deeply to the exonic circRNAs characteristics, approximately the 23 % of the predicted circRNAs contained 2 exons, followed decreasingly by forms with 3, 4, 5, 1 and 6 exons (Figure 1B). Exonic circRNAs with greater length accounted for only 10 % of the total hits (Figure 1B). Of note, solely one of the predicted exonic circRNAs, the one derived from the Vacuolar Protein Sorting 13 Homolog A (*VPS13A*) gene, contained up to 21 exons. In addition, around 60 % of genes only showed one exonic circRNA form (that is, one isoform), while the 21 % had two different exonic circRNAs isoforms per gene (Figure 1C). Remarkably, the highest number of different circular isoforms derived from a single gene was found for the Eukaryotic Translation Initiation Factor 4 Gamma 3 (*EIF4G3*) gene, with up to 13 different predicted circRNAs isoforms.

The Gene Ontology (GO) enrichment analyses suggested the involvement of the predicted sperm exonic circRNAs in biological processes such as chromosome organization, chromatin organization, chromatin remodeling and regulation of gene expression, among others (Table 1; Figure 1D). Likewise, RNA binding, chromatin binding, histone binding and protein binding were enriched in the molecular function analysis, while the terms chromosome and nucleus were found enriched in the cellular component analysis (Table 1; Figure 1D). Noteworthy, 57 out of the 737 exonic circRNAs in common between biological replicates derived from genes functionally involved in epigenetic regulation, such as chromatin remodeling factors, histone modification factors and transcription factors (Supplementary Table 5).

Selected circRNAs candidates for experimental validation

Twelve sperm exonic circRNAs predicted by the KNIFE algorithm were selected for experimental validation. For this selection, different factors, such as the abundance of the predicted circRNA junction calculated by the KNIFE algorithm number of reads, as well as the biological function of the gene from which the circRNA is arisen, were taken into consideration. The twelve selected exonic circRNAs derived from 10 genes with the

Table 1. Gene ontology (GO) terms enriched on the dataset of 737 sperm exonic circRNAs (n=3). GO names are listed together with the gene count and the Bonferroni adjusted p-value. Processes related to epigenetic functions are highlighted in bold. Selection of the relevant GO terms according to sperm function has been included**.**

Biological process GO terms	Gene Count	P-value Bonferroni adjusted
organelle organization (GO:0006996)	169	7,11E-13
regulation of macromolecule metabolic process (GO:0060255)	234	5,39E-07
chromosome organization (GO:0051276)	**65**	**4,32E-06**
cellular component organization (GO:0016043)	215	2,10E-06
chromatin organization (GO:0006325)	**49**	**1,15E-05**
regulation of gene expression (GO:0010468)	**174**	**1,76E-04**
sexual reproduction (GO:0019953)	51	3,35E-04
regulation of RNA metabolic process (GO:0051252)	149	2,07E-03
multicellular organismal reproductive process (GO:0048609)	49	2,08E-03
reproductive process (GO:0022414)	72	1,27E-03
reproduction (GO:0000003)	72	1,34E-03
developmental process involved in reproduction (GO:0003006)	54	2,68E-03
chromatin remodeling (GO:0006338)	**19**	**3,06E-03**
male gamete generation (GO:0048232)	37	7,22E-03
spermatogenesis (GO:0007283)	36	9,56E-03
multi-organism reproductive process (GO:0044703)	53	1,38E-02
regulation of DNA repair (GO:0006282)	15	3,68E-02
Molecular function GO terms	**Gene Count**	**P-value Bonferroni adjusted**
protein binding (GO:0005515)	**388**	**2,09E-07**
binding (GO:0005488)	**454**	**3,65E-05**
histone binding (GO:0042393)	**20**	**3,15E-03**
RNA polymerase II-specific DNA-binding transcription factor bin	**24**	**4,08E-03**
chromatin binding (GO:0003682)	**36**	**5,57E-03**
G protein-coupled receptor activity (GO:0004930)	4	7,97E-03
ubiquitin-like protein binding (GO:0032182)	13	1,27E-02
enzyme binding (GO:0019899)	94	2,01E-02
RNA binding (GO:0003723)	**75**	**2,07E-02**
Cellular component GO terms	**Gene Count**	**P-value Bonferroni adjusted**
DNA repair complex (GO:1990391)	**45**	**0,00E+00**
centriole (GO:0005814)	181	9,66E-12
microtubule organizing center (GO:0005815)	276	3,53E-09
microtubule cytoskeleton (GO:0015630)	432	4,26E-09
catalytic complex (GO:1902494)	194	1,88E-08
nucleoplasm (GO:0005654)	**198**	**7,31E-08**
chromosome (GO:0005694)	**349**	**3,24E-05**
plasma membrane bounded cell projection (GO:0120025)	213	4,80E-05
nuclear lumen (GO:0031981)	213	4,80E-05
non-membrane-bounded organelle (GO:0043228)	213	4,80E-05
intracellular non-membrane-bounded organelle (GO:0043232)	43	9,36E-05
nucleus (GO:0005634)	**100**	**5,04E-04**

followed characteristics: 6 circRNAs derived from 5 genes involved in transcription regulation (*ZNF608, BPTF, NFATC3, ANKRD12* and *TOP1*); 2 circRNAs derived from 2 genes participating in chromatin remodeling and histone modification (*BAZ1A* and *VRK1*); 3 circRNAs derived from 2 genes with roles in other processes such as cell growth and centrosome formation (*TBC1D1* and *ALMS1*); and, finally, 1 circRNA derived from the *LINC02050* gene, which is the most abundant circRNA in our dataset. The

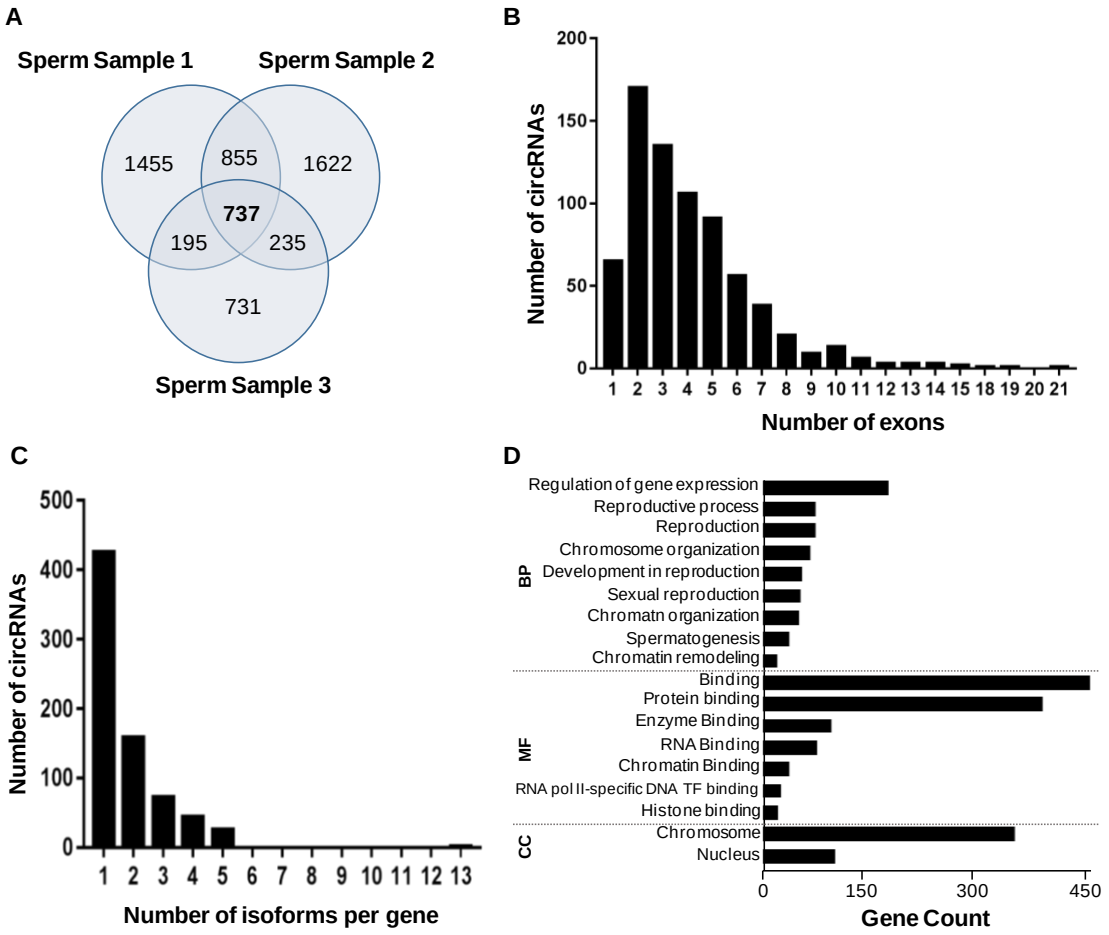


Figure 1. Global analysis of the predicted sperm exonic circRNAs. (A) Total number of exonic circRNAs predicted by the KNIFE algorithm in each of the three human sperm samples included in this study. The overlap between biological replicates is highlighted in bold and has been used for further analysis. (B) Number of exons contained in each predicted sperm circRNA. Only results corresponding to the 737 exonic circRNAs in common between the three biological replicates are shown. (C) Number of isoforms found for the genes corresponding to the predicted circRNAs in common between biological samples. (D) Selection of the relevant GO terms according to reproduction and epigenetics. Grey discontinuous arrows separate biological process (BP), molecular function (MF) and cellular component (CC).

functional involvement of the genes was confirmed using information contained in the GeneCards® database (https://www.genecards.org/) (Stelzer et al., 2016). The specific characteristics of each exonic circRNA candidate are indicated in Table 2. The length of the selected sperm exonic circRNAs candidates was variable, with an average of 1,350 bp and ranging from 250 bp, corresponding to the *LINC02050* derived circRNA, to 6,336 bp, corresponding to the *ALMS1* exon 7 to exon 9 derived circRNA (Table 2). The circRNAs derived from the *ZNF608*, *TBC1D1*, *ANKRD12*, and *VRK1* genes contained extremely large intronic flanking regions, showing an average length of almost 34,000 bp (Table 2). On the contrary, the exonic circRNAs derived from the *LINC02050* and *BPTF* exon 21 to exon 27 genes contained short intronic flanking regions of less than

3,000 bp. The remaining exonic circRNAs candidates contained medium intronic flanking regions, with an average length of 12,000 bp (Table 2).

Regarding the selection of intronic circRNAs for experimental validation, the application of the strict criteria described in the methods section resulted in 4 sperm IEs candidates derived from the genes *QRICH1*, *AP2A1*, *ZMYND15* and *TRIM66* (Table 2; Supplementary Table 2; Supplementary Figure 3). These genes are functionally involved in the regulation of RNA polymerase II transcription, in the regulation of vesicle transport by the plasma membrane, in the regulation of histone deacetylases and spermiogenesis, and in transcriptional repression, respectively. Remarkably, the total intronic length of the IEs was short, with an average length of 350 bp.

Determination of the circularization junction of potential circRNAs

PCR amplification utilizing specific divergent primers followed by direct sequencing confirmed the presence of circularized junctions for all the exonic and IEs circRNAs candidates selected for experimental validation, thus corroborating their circular form (Supplementary Figure 4; Supplementary Figure 5).

RNase R treatment optimization

Since no consensus exists on the amount of RNase R required to specifically digest linear RNAs (Salzman et al., 2012, 2013; Memczak et al., 2013; Dong et al., 2016), we conducted an optimization test using high-quality RNA from somatic cells and spermatozoa (Supplementary Figure 6). When using 0.5 U of RNase R per 100 ng of sperm RNA, a 4-fold decrease in the amplification of the linear *PRM1* using convergent primers, and a small increase in the amplification of the IE circRNA located in the *TRIM66* gene with divergent primers were observed (Supplementary Figure 6). This behavior mimicked the results obtained when 5 U of RNase R were used in somatic HEK cells for the linear and circular isoforms of the *CAMSAP1* gene (Supplementary Figure 6). In particular, the linear isoforms amplified by convergent primers were degraded after RNase R treatment, while circular isoforms amplified by convergent and divergent primers were resistant to the treatment. Based on these results, 0.5 U of RNase R was set as the optimal condition to treat sperm RNA. In addition, the amplification of the *PRM1* transcript with convergent primers and of the IE *TRIM66* with divergent primers were employed in further experiments as quality control to test the RNase R treatment sensitivity and resistance, respectively.

Table 2. Characteristics of the predicted sperm circRNAs experimentally validated in this study.

Gene Name	Genome h38 circular junction position	Transcript ID	Circular junction	Average reads	circRNA length (bp)	flanking intron 5' length (bp)	flanking intron 3' length (bp)	flanking exon 5' length (bp)	flanking exon 3' length (bp)
Exonic circRNAs									
ALMS1	chr2:73455295-73432198	ENST00000613296.5	Exon 9 - Exon 7	165,0	6336	5644	34338		
	chr2:73559142-73557220	ENST00000613296.5	Exon 15 - Exon 14	104,7	306	6782	13119		
LINC02050	chr3:80788969-80784041	ENST00000473938.1	Exon 4 - Exon 3	5543,3	250	3795	267		
TBC1D1	chr4:38103157-38089932	ENST00000261439.9	Exon 15 - Exon 13	51,0	507	35593	12552		
ZNF608	chr5:124701014-124701269	ENST00000306315.9	Exon 2 - Exon 2	325,3	256	42814	51316		
BAZ1A	chr14:34792775-34862322	ENST00000358716.8	Exon 12 - Exon 3	874,0	1397	12169	6773		
VRK1	chr14:96860735-96833467	ENST00000216639.8	Exon 11 - Exon 2	87,7	1073	36019	15294		
NFATC3	chr16:68123121-68121987	ENST00000329524.8	Exon 2- Exon 2	595,7	1135	36202	3326		
BPTF	chr17:67975958-67945409	ENST00000306378.11	Exon 27 - Exon 21	592,7	2026	1036	6293		
	chr17:67948306-67945409	ENST00000306378.11	Exon 23 - Exon 21	174,0	1226	1036	11234		
TOP1	chr20:41101353-41092472	ENST00000361337.3	Exon 13 - Exon 9	101,7	694	7903	11428		
ANKRD12	chr18:9221999-9182382	ENST00000262126.9	Exon 8 - Exon 2	286,3	994	45416	32211		
Intronic element (IE) circRNAs									
QRICH1	chr3:49032819-49033118	ENST00000395443.7	Partial intron 7 - Partial intron 7		299			109	152
AP2A1	chr19:49806256-49806593	ENST00000354293.9	Partian intron 22 - Partial intron 22		337			135	432
ZMYND15	chr17:4743462-4743701	ENST00000269289.10	Partial intron 6 - Partial intron 6		239			153	81
TRIM66	chr11:8621345-8621644	ENST00000299550.10	Partial intron 14 - Partial intron 14		299			175	290

Sperm circRNAs show three different resistance patterns to the RNase R treatment

The RNase R treatment, conducted to validate the resistance of the selected exonic and IEs circRNAs in sperm RNA samples, revealed three different RNase R resistance patterns (Figure 2). The exonic circRNAs derived from the *LINC02050* and *ZNF608* genes and the IEs circRNAs derived from the *ZMYND15*, *AP2A1* and *TRIM66* genes were classified as RNase R resistance pattern 1. This group of circRNAs showed resistance to the RNase R treatment for both convergent and divergent primers amplifications, indicating that the RNA arising from those exonic and intronic regions were mainly circRNAs in the mature sperm (Figure 2A). RNase R resistance pattern 2 was observed for the *ALMS1* exon 14 – exon 15 derived circRNA. This pattern mimicked the behavior of circRNAs found in somatic cells, where both the circRNAs and their linear cognates co-exist (Figure 2B; Supplementary Figure 6A). In this way, the resistance to the RNase treatment was only observed using divergent primers, since only the circRNA isoform was targeted. However, when convergent primers were used, the quantity of RNA after the RNase R treatment decreased, due to the degradation of most of the linear form and the amplification of only the circular form (Figure 2B). Of note, the use of divergent primers for the *ALMS1* exon 14 – exon 15 derived circRNA showed a slight increase of amplification, which was probably due to a higher efficiency of the circRNA retrotranscription after the degradation of the major part of the linear RNA. Finally, the RNase R resistance pattern 3 was observed in the exonic circRNAs derived from the *VRK1* and *NFATC3* genes, which showed a high sensitivity to the RNase R treatment after both convergent and divergent primers amplifications. This behavior suggested the decircularitzation of the targeted circRNAs (Figure 2C).

To further test whether these different RNase R resistance patterns of circRNAs were exclusive of the mature transcriptionally and translationally inert spermatozoa or could be also found in immature male germ cells, we tested the RNase R resistance of the same exonic and IE circRNAs in testicular tissue (Figure 3). Only the circRNAs derived from the *LINC02050, TRIM66* and *ZMYND15* genes showed similar behavior than the RNase R resistance pattern 1 observed in the sperm samples (Figure 3A). On the contrary, the circRNAs derived from the *ALMS1*, *NFATC3*, *ZNF608*, *VRK1* and *AP2A1* genes followed the RNase R resistance pattern 2 (Figure 3B). Of note, none of the tested circRNAs showed the clear decircularitzation corresponding to the RNase R resistance pattern 3 observed in mature sperm.

The resistance to the RNase R treatment for the other sperm circRNAs candidates derived from the genes *BAZ1A*, *BPTF*, *ANKRD12*, *TBC1D1*, *TOP1*, *ALMS1* exon 7 – exon 9 and *QRICH1* was not reliably established, due to the high Ct values obtained

after PCR amplification (Ct>30; data not shown), suggesting a low abundance of these circRNAs or an inefficient PCR.

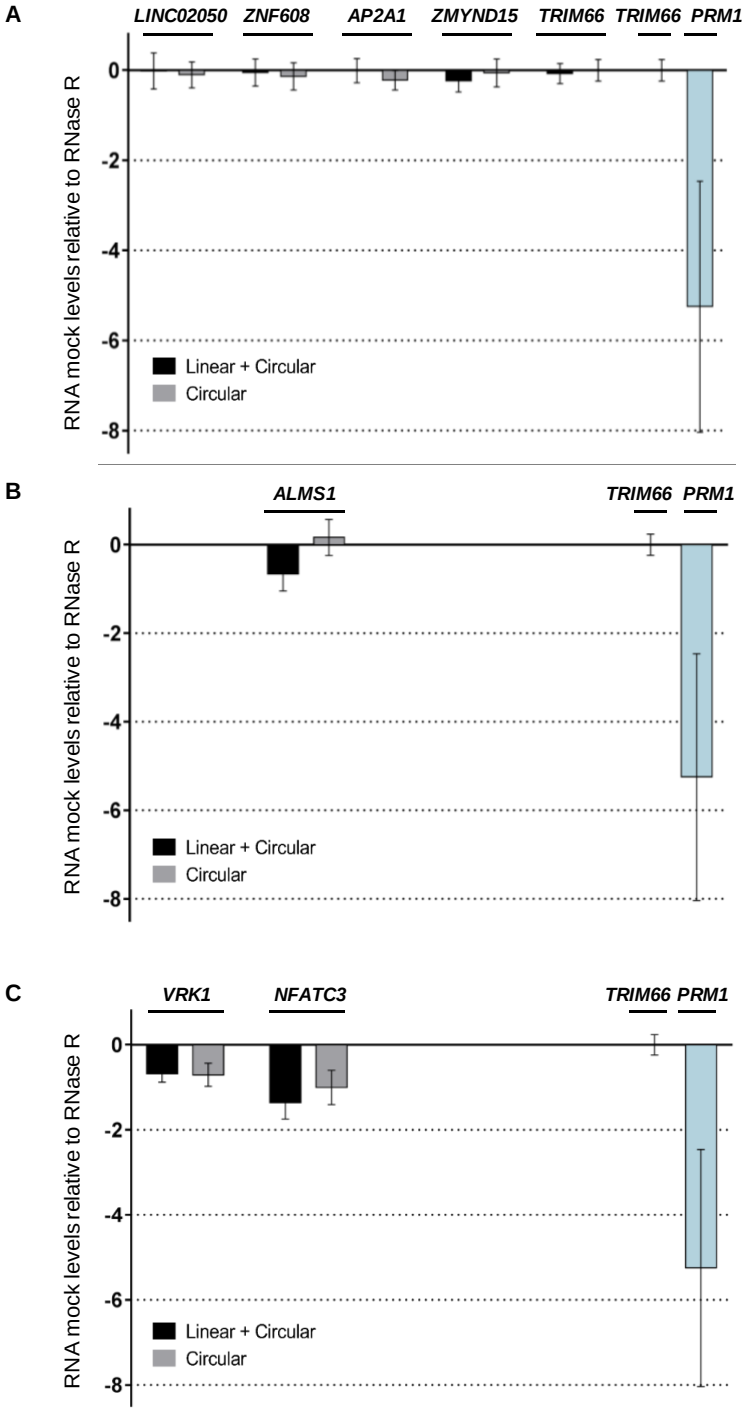


Figure 2. Validation of the sperm circRNAs resistance to the RNase R treatment (0.5 U). (A) circRNAs presented majorly in its circular form in the sperm cell (RNase R resistance pattern 1). (B) circRNAs with typical dual somatic pattern (RNase R resistance pattern 2). (C) circRNAs sensible to the RNase R treatment (RNase R resistance pattern 3). Results corresponding to the assessment of the *TRIM66* gene with divergent primers, used as positive control of circularization, and of the *PRM1* gene with convergent primers, used as a control of the RNase R digestion efficiency, are showed in light blue.

The circRNA derived from the *VRK1* gene shows a specific cleavage causing its decircularization

The decircularitzation observed for some sperm circRNAs, corresponding to the RNase R resistance pattern 3, could be due to either random circRNA degradation or a specific cleavage (Figure 2C). To test both hypotheses, we deeply analyzed the exonic circRNA derived from the *VRK1* gene by designing different combinations of divergent and convergent primers across the 10 exons comprised in the circRNA (see material and

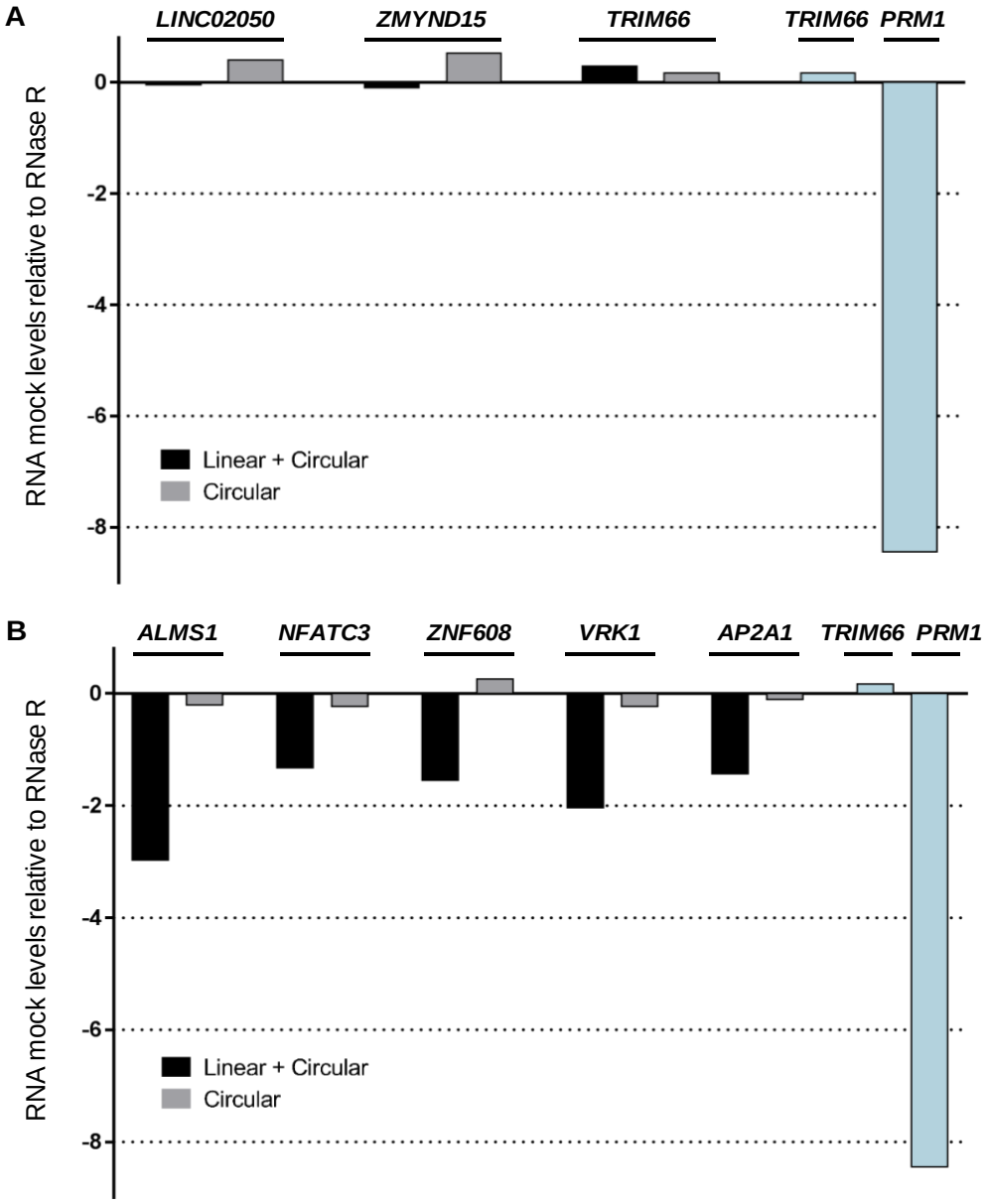


Figure 3. Validation of the circRNAs resistance to the RNase R treatment with testis tissue. (A) circRNAs presented majorly in its circular form (RNase R resistance pattern 1). (B) circRNAs with typical dual somatic pattern (RNase R resistance pattern 2). Results corresponding to the assessment of the *TRIM66* gene with divergent primers, used as positive control of circularization, and of the *PRM1* gene with convergent primers, used as a control of the RNase R digestion efficiency, are showed in light blue.

methods section; supplementary figure 2A). Of note, when studying the combination of the divergent primers sets E7 – E2, E8 – E2, E9 – E2 and E10 – E2, a minor form of an additional circRNA isoform, which derived from the circular junction of exon 10 to exon

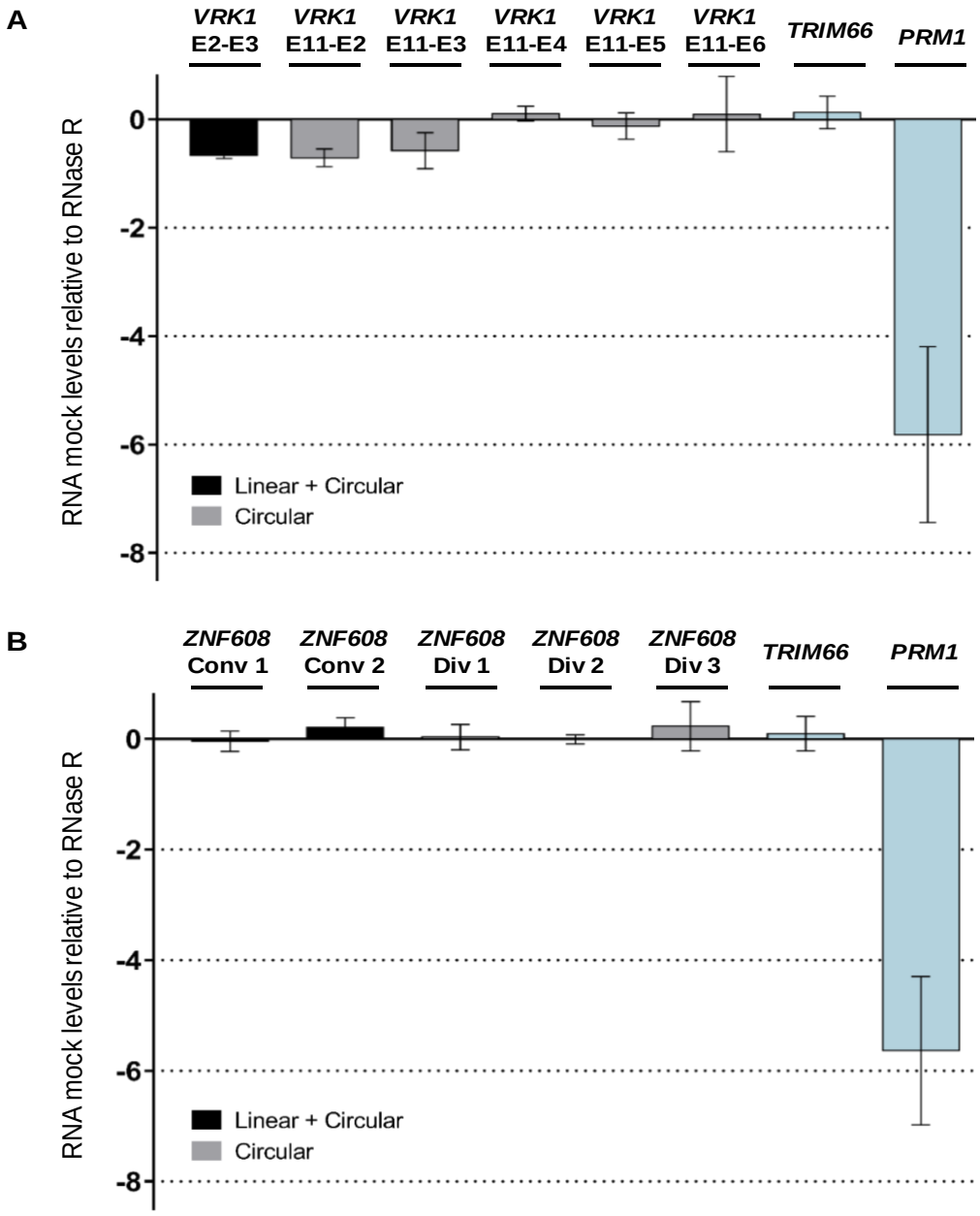


Figure 4. Resistance to RNase R treatment of the amplicons obtained when using different combinations of convergent and divergent primers. (A) Resistance to RNase R treatment of the circRNA derived from the *VRK1* gene using different combinations of convergent and divergent primers (n = 5). (B) Resistance to RNase R treatment of the circRNA derived from the *ZNF608* gene using different combinations of convergent and divergent primers (n = 4)).

2 of the *VRK1* gene, was detected (Supplementary Figure 7). Therefore, all those divergent primer sets whose amplification product included the E10 – E2 junction of the

VRK1 gene were excluded from the analysis (Supplementary Table 3; Supplementary Figure 2). The products obtained with the convergent primers E2 – E3 and the divergent primers sets E11 – E2 and E11 – E3 showed lower resistance to the RNase R treatment than the products obtained with the divergent primer combinations E11 – E4, E11 – E5 and E11 – E6 (Supplementary Table 3; Figure 4A). These results suggested the existence of a specific cleavage site between the exons 3 and 4, which would result in the decircularization of this circRNA (Figure 5) and, thus, in the generation of a linear RNA containing the E11 – E2 junction that is degraded by the RNase R, due to the presence of a free 3'-end. In order to assess the efficiency of this methodology, we employed the circRNA derived from *ZNF608,* which showed resistance to the RNase R treatment (RNase R resistance pattern 1), to verify its circularity. In contrast to the results observed for the *VRK1* circRNA, all the amplicons of the circRNA derived from *ZNF608* that were obtained after using different combinations of divergent primers were resistant to the RNase R treatment. These results suggested the absence of any specific cleavage site, and demonstrated that the *ZNF608* derived exonic circRNA was intact in mature spermatozoa (Supplementary Table 3; Figure 4B).

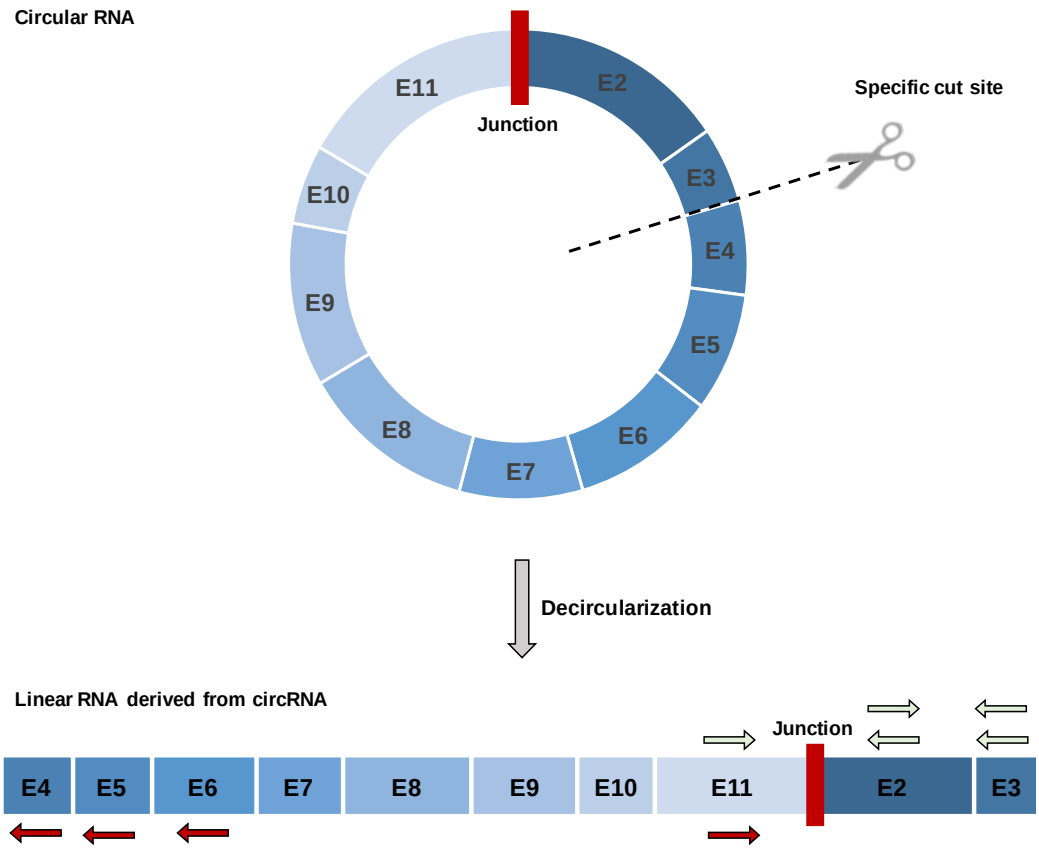


Figure 5. Schematization of the circRNA derived from *VRK1* gene and its specific region of decircularization (upper part) and the resultant linear form after specific cut region (lower part). Top green arrows from the linear form symbolize primer sets that amplify and lower red arrows symbolize primer sets that do not amplify.

Potential roles of sperm circRNAs on fertility, sperm functionality, early embryo development and beyond

To further analyze the possible contribution of sperm exonic circRNAs to male fertility, we compared the 737 exonic circRNAs predicted in the three samples analyzed in this study with the 648 SREs proposed as necessary for the natural conception in a previous publication (Jodar et al., 2015). Interestingly, we found that 44 SREs critical to achieving a live birth deriving from 25 different genes were embedded inside 49 predicted sperm circRNAs (Supplementary Table 6). From these, the *BPTF*, the SWI/SNF Related, Matrix Associated, Actin Dependent Regulator of Chromatin Subfamily C Member 1 (*SMARCC1*) and the DNA Methyltransferase 1 (*DNMT1*) genes have been described as involved in chromatin remodeling and DNA modification (Supplementary Table 5). Additionally, 2 out of these 44 SREs embedded inside exonic circRNAs were found altered in infertile idiopathic couples unable to naturally conceive (Jodar et al., 2015), namely the Serine/Threonine Kinase 39 (*STK39*) and the G Protein Subunit Beta 1 (*GNB1*; Supplementary Table 6). To gain insight also into the possible participation of sperm circRNAs in the sperm proper functionality, we compared the 737 exonic circRNAs described herein with the list of 487 SREs associated with obesity reported in Swanson et al., 2020 Eight SREs linked to obesity and epigenetics, which derived from 7 different genes, were found embedded inside 9 exonic circRNAs from the sperm (Supplementary Table 6). Remarkably, two of them, namely NIPBL Cohesin Loading Factor (*NIPBL*) and *DNMT1* (Supplementary Table 6), have been associated with epigenetic functions as histone modification eraser cofactor and DNA modification factor, respectively (Supplementary Table 5).

In addition, we explored whether circRNAs could play a role in early embryogenesis by determining which sperm circRNAs were also present in the zygote and absent in the oocyte. Of note, we identified 287 and 114 exonic circRNAs in the oocyte and the zygote, respectively (Figure 6). This integrative analysis evidenced that 9 exonic circRNAs in the human zygote could be exclusively delivered by the sperm, since were not detected in the oocyte (Figure 6). These potential paternally-derived embryonic circRNAs were associated to the genes GLE1 RNA Export Mediator (*GLE1*), SHOC2 Leucine Rich Repeat Scaffold Protein (*SHOC2*), Mitochondrial Ribosome Associated GTPase 1 (*MTG1*), Centrosome and Spindle Pole Associated Protein 1 (*CSPP1*), Kinesin Family Member 20B (*KIF20B*), ELAV Like RNA Binding Protein 2 (*ELAVL2*), Bromodomain Adjacent to Zinc Finger Domain 1B (*BAZ1B*), Erythrocyte Membrane Protein Band 4.1 Like 2 (*EPB41L2*) and *VRK1* (Table 3). According to the GeneCards® database (https://www.genecards.org/) (Stelzer et al., 2016), these paternally-derived embryonic circRNAs seemed to be associated to epigenetic roles in the zygote, such as *BAZ1B* and

VRK1, which are modification writer factors that act phosphorylating histones (Supplementary Table 5).

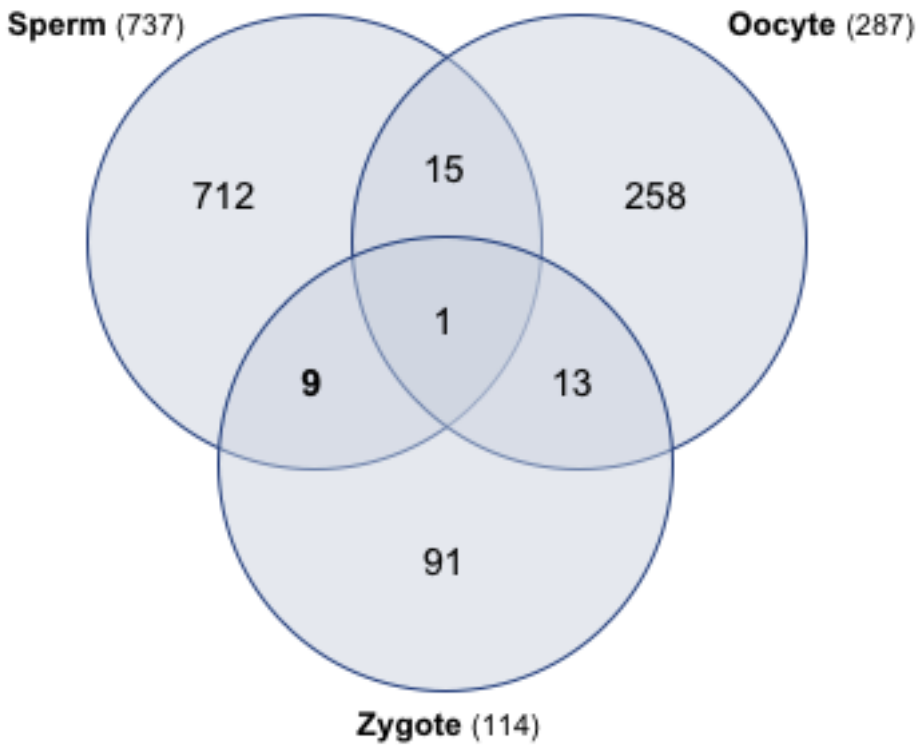


Figure 6. Overlapping of the exonic circRNAs corresponding to sperm, oocyte and zygote according to KNIFE algorithm. We employed sperm exonic circRNAs in common between the three biological replicates and exonic circRNAs from single-cell oocytes and single-cell zygotes datasets from Dang et al., 2016. Oocyte technical and biological replicates were treated as separate replicates to strictly discard a maternal origin. Only exonic circRNAs in common in all zygote biological and technical replicates were included. Only exonic predicted circRNAs with ≥ 2 reads and *p* value score ≥ 0.9 were considered for the analysis.

Potential mechanisms underlying sperm circRNA functions

It has been proposed that circRNAs could act as sponges of miRNAs and proteins, as enhancers or coordinators of proteins, mRNAs and DNAs, and as templates for translation. In order to check the ability of sperm circRNAs to act as miRNA sponges, we assessed the presence of repetitive miRNA seeds in the validated exonic and IE circRNAs, as well as in the paternally-derived embryonic circRNAs. From those, the majority of the tested sperm circRNAs contained 2 or 3 7-mer repetitions, and only the paternally-derived embryonic circRNA derived from the *EPB41L2* gene contained 4 7-mer repetitions (Supplementary Table 7). These results suggested that miRNA sponging might not probably be the principal function of sperm circRNAs.

Additionally, to gain more insight into the sperm circRNA translatable potential to functional peptides, we predicted potential ORFs for the validated exonic and IE

Table 3. Characteristics of the paternally-derived embryonic circRNAs.

Gene Name	Genome h38 circular junction position	Transcript ID	Circular junction	Average sperm reads	Average zygote reads	Functions*
EPB41L2	chr6:130955105-130956499	ENST00000337057.8	Exon 3 - Exon 2	216,7	5,8	Role in the cell-cycle anaphase
BAZ1B	chr7:73459536-73470483	ENST00000339594.9	Exon 13 - Exon 8	8,3	3,0	Atypical tyrosine-protein kinase that plays a central role in chromatin remodeling and acts as a transcription regulator
CSPP1	chr8:67137603-67131951	ENST00000262210.9	Exon 15 - Exon 14	70,7	4,8	Role in cell-cycle progression, spindle organization and regulates cytokinesis
GLE1	chr9:128515639-128508876	ENST00000309971.8	Exon 3 - Exon 2	514,0	15,5	Required for export mRNAs containing poly(A) tails from the nucleus into the cytoplasm
ELAVL2	chr9:23762006-23765099	ENST00000380110.8	Exon 3 - Exon 2	87,0	2,3	RNA-binding protein that binds to the 3' untranslated region (3'UTR) of target mRNAs
KIF20B	chr10:89762835-89751346	ENST00000260753.8	Exon 29 - Exon 24	54,7	13,5	Required for completion of cytokinesis and related to ATPase activity and microtubule motor activity
SHOC2	chr10:110985765-110964125	ENST00000265277.9	Exon 3 - Exon 2	82,3	8,3	Scaffold in the activation of the Ras and MAP Kinase signaling pathways
MTG1	chr10:133402773-133395713	ENST00000317502.11	Exon 9 - Exon 2	18,3	6,0	Plays a role in the regulation of the mitochondrial ribosome assembly and of translational activity
VRK1	chr14:96860735-96833467	ENST00000216639.8	Exon 11 - Exon 2	87,7	8,8	Kinase that participate in chromatin remodeling and histone modifications

* According to GeneCards® database (https://www.genecards.org/)

circRNAs, as well as for the paternally-derived embryonic circRNAs predicted herein, and we provided the potential resulting polypeptides (Supplementary Table 8). Excluding the paternally-derived embryonic circRNA arisen from the *CSPP1* gene, in which no ORFs were detected, all the other tested circRNAs seemed to possess at least 1 ORF that would give rise to a functional peptide. Surprisingly, the results suggested that paternally-derived embryonic circRNAs may produce truncated forms of the corresponding canonical proteins, being some of them coding specifically for their functional domains (Supplementary Table 8; Supplementary Figure 8). For instance, a predicted ORF for the *MTG1* derived circRNA was found associated with the translation of the catalytic domain of its corresponding protein, which is a circularly permuted-type guanine nucleotide-binding (CP-type G) domain (Supplementary Figure 8A). Likewise, a predicted ORF for the *VRK1* derived circRNA was identified as being able to translate the protein kinase domain of the corresponding protein (Supplementary Figure 8B). Altogether, our results strongly support the hypothesis that sperm circRNAs could act as potential protein-coding RNAs.

DISCUSSION

In the present study, we have reported and validated the presence of a complex population of circRNAs contained in the mature sperm. By applying the reliable KNIFE algorithm (Szabo et al., 2015), we have identified up to 5,830 different exonic circRNAs in high-quality sperm RNA samples, from which 737 were found in common in the three biological replicates analyzed. In addition, we have reported, for the first time, that the most abundant IEs previously identified in Jodar et al., 2015 (Jodar et al., 2015) are in fact intronic circRNAs.

The analysis of the resistance to the RNase R treatment of the predicted circRNAs, not previously conducted for the male gamete, has allowed us to unravel the main peculiarities associated to human sperm circRNAs. As a differential feature with somatic cells in which linear and circular forms derived from the same gene co-exist (RNase R resistance pattern 2 described herein) (Salzman et al., 2012, 2013; Barrett et al., 2015), we have observed that many RNAs from the sperm cell are exclusively present in their circular form (RNase R resistance pattern 1). In addition, since the linear cognates of these circRNAs were detected in the human testis, we propose a gradual degradation of linear forms taking place during spermatogenesis, resulting in a selective retention of circular forms in the mature sperm, similarly to what has been observed in mice (Tang et al., 2020). Intact sperm circRNAs that are only detected in their circular form, such as the IE circRNA *TRIM66*, would be good RNA integrity markers for sperm RNA. Their sensitivity to the treatment would allow determining the poor quality of a sperm RNA preparation,

overcoming the lack of sperm integrity quality controls similarly to the wide use of RIN value in somatic cells (Schroeder et al., 2006), but further experiments are required to implement it. However, not all the circRNAs seem to be maintained intact in the male gamete, since some of them showed sensitivity to the RNase R treatment (RNase R resistance pattern 3) in all the samples assessed. Interestingly, a deeper analysis revealed a non-random decircularization of those circRNAs in the mature sperm, which leaded us to hypothesize that some sperm circRNAs could co-exist with their circRNA turnover machinery. Although circRNA turnover mechanisms have been poorly studied, two different pathways have been proposed to take part in it. One of the suggested mechanisms proposes miRNAs as responsible of triggering circRNA specific cleavage through the argonaute 2 (AGO2), as observed for the miR-671 binding to the circRNA ciRS-7 (Hansen et al., 2011; Kleaveland et al., 2018). Alternatively, it has been proposed that circRNAs containing the RNA modification m6A might be subjected to endoribonuclease cleavage by the complex ribonuclease P (RNase P) – multidrug resistance-associated protein 1 (MRP), together with the m6A reader YTH domain-containing family protein 2 (YTHDF2) and the adaptor protein heat-responsive protein 12 (HRSP12) (Park et al., 2019). Remarkably, the spermatozoa contain the functional proteins YTHDF2 and HRSP12, as well as two proteins from the RNase P – MRP complex, according to the compiled human sperm proteome (Castillo et al., 2018). Therefore, the presence in the spermatozoa of different components responsible for regulating the circRNAs turnover, either specific miRNAs or endoribonucleases, suggests a selective cleavage of some sperm circRNAs, which could occur during gamete maturation or during the transit of sperm from the testis to the oocyte. However, the exact mechanism remains to be elucidated.

Sperm circRNAs could be remnants from spermatogenesis and sperm maturation, making them potentially useful as biomarkers of male fertility. Interestingly, we found SREs linked to achieving a live birth by natural conception, reported in Jodar et al., 2015 (Jodar et al., 2015), embedded in the sperm circRNAs reported in this study. In addition, *STK39* and *GNB1* SREs, which were found absent in some patients with idiopathic infertility unable to achieve a live birth, have been found embedded in two predicted sperm circRNAs (Jodar et al., 2015). Therefore, a probable role of sperm circRNAs on the establishment of the fertility state is suggested (Chioccarelli et al., 2019; Gòdia et al., 2020; Tang et al., 2020). Our results are supported by a recent study describing the presence of differentially abundant circRNAs between high-quality and poor-quality spermatozoa according to motility and morphology parameters, which have an impact on the fertility ability of the male gamete (Chioccarelli et al., 2019). Likewise, an increase in the number of sperm circRNAs in fertile donors in comparison with infertile men has been reported, reinforcing

the idea that a deregulation in the circRNA cargo of spermatozoa could lead to infertility (Tang et al., 2020). Our results have also showed that some SREs associated with increased BMI (Swanson et al., 2020) are embedded inside sperm circRNAs, such as the NIPBL Cohesin Loading Factor (*NIPBL*), the Leucine Rich Repeats And IQ Motif Containing 1 (*LRRIQ1*), the UTP18 Small Subunit Processome Component (*UTP18*), the SRSF Protein Kinase 2 (*SRPK2*), the Early Endosome Antigen 1 (*EEA1*), the *DNMT1* and the Eukaryotic Translation Initiation Factor 3 Subunit A (*EIF3A*) SREs (Swanson et al., 2020). This result is highly interesting due to the fact that sperm RNAs found deregulated due to life-style habits, such as obesity, have been associated to altered sperm functions (Donkin et al., 2016; Swanson et al., 2020).

Growing evidence show that specific sperm RNAs could act as epigenetic marks and regulate embryonic gene expression beyond fertilization (Ostermeier et al., 2004; Krawetz et al., 2011; Sendler et al., 2013; Chen et al., 2016; Donkin et al., 2016; Sharma et al., 2016; Gòdia et al., 2018; Jodar, 2019). The remarkable stability of circRNAs, in conjunction with its association with GO terms related to chromatin and epigenetic functions in sperm circRNAs datasets (Chioccarelli et al., 2019, 2020; Ragusa et al., 2019; Gòdia et al., 2020; Tang et al., 2020), point out the sperm circRNAs as good candidates to regulate embryo gene expression. For instance, the most abundant sperm circRNA, the *LINC02050* derived circRNA would probably be involved in gene expression, since lncRNAs are usually implicated, in processes such as chromatin remodeling, mRNA translation and transcriptional activation, among others (Zhang et al., 2019). However, this hypothesis requires further validation. Other sperm circRNAs that might be involved in gene expression regulation are those derived from the *NFATC3* gene. *NFATC3* belongs to the NFAT transcription factors family, which have been associated with the regulation of the differentiation lineage during embryo development (Li et al., 2011; Swanson et al., 2020). Additionally, the fact that some IEs are specifically retained in the mature sperm in their circular form, showing higher stability, suggests that these intronic circRNAs are more than just lariat RNAs. These results are supported in other cell types by studies identifying, majorly in the nucleus, intronic circRNAs escaping debranching. These intronic circRNAs hold regulatory functions, such as being associated with RNA polymerase II and regulate the expression of their parental genes (Zhang et al., 2013; Panda et al., 2017).

In order to identify sperm circRNAs that could be essential for the early embryo, we assessed which embryonic circRNAs might be provided exclusively through the paternal lineage, and found nine paternally-derived embryonic circRNAs candidates. Remarkably, one paternally-derived embryonic circRNAs is the exonic circRNA derived from the *GLE1* gene, which is one of the ten most abundant circRNAs in sperm and also contain one of the 648 SRE required to achieve live birth by natural conception (Jodar et al., 2015). The

relevance of *GLE1* gene in embryogenesis relies on the embryonic lethally observed in zebrafish knock-out *GLE1* embryos due to neural development deficiencies (Seytanoglu et al., 2016). Surprisingly, the exonic circRNA derived from *VRK1* gene, for which we have demonstrated a specific non-stochastic decircularization, was also defined as a potential paternally-derived embryonic circRNA. Functional VRK1 is a kinase that phosphorylates histones and its RNA circular form has been identified also in other cell types (Yan et al., 2017). In addition, VRK1 stimulates the telomerase activity and its deficiency in mice results in germ cells apoptosis due to shorter telomeres and abnormal telomere arrangement (Choi et al., 2012). Since telomere length is important for embryo development (Turner et al., 2010; Jodar et al., 2020), it could be possible that the exonic circRNA derived from the *VRK1* gene harbored a role associated with telomere length maintenance in the developing embryo.

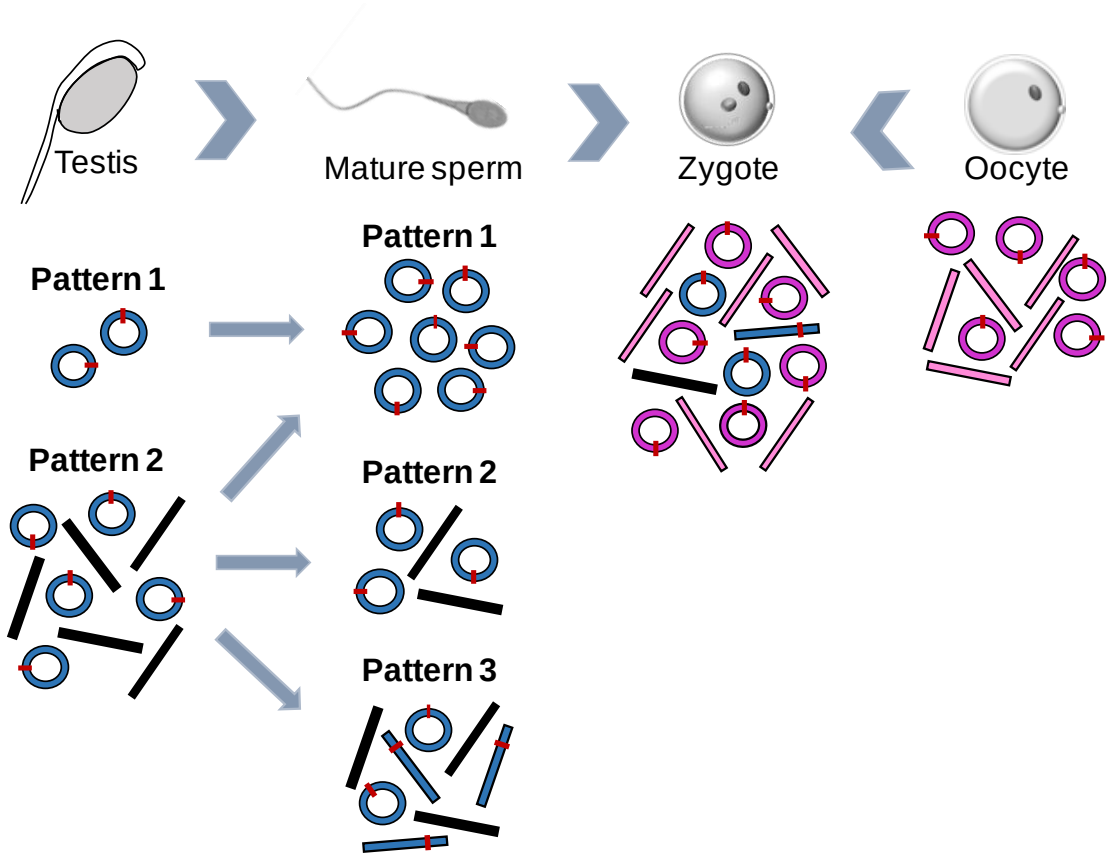


Figure 7. Overall of the different RNase R resistance patterns found in testis and sperm circRNAs and the potential contribution to the embryo during fertilization, together with oocyte potential circRNA cargo. Paternal cargo is highlighted in blue circles (circRNAs), black rectangles (linear forms) and blue rectangles (decircularized circRNAs). Maternal cargo is highlighted in pink circles (circRNAs) and light pink rectangles (linear forms). Red rectangles indicate the circularization junction.

Inferring the potential functions of sperm circRNAs is crucial to gain knowledge in their effective involvement. The most popular function associated with circRNAs is the miRNA

sponging (Hansen et al., 2013; Barrett and Salzman, 2016; Greene et al., 2017). However, according to our results, it does not seem to be the principal role of sperm circRNAs. Other recent proposed actions for circRNAs are the regulation of gene expression and their translation into functional peptides (Wang and Wang, 2015; Legnini et al., 2017; Yang et al., 2017). Indeed, a recent study has postulated that sperm circRNAs are potentially translated (Ragusa et al., 2019), and Tang and colleagues have demonstrated that the presence of large intact ORFs and m^6A-modified start codons at the junction sites of circRNAs in male germ cells are translatable to peptides (Tang et al., 2020). Our data supports this hypothesis, since the sperm circRNAs experimentally validated in this study showed at least one predicted ORF that could be potentially translated to a polypeptide. It has been described that, during spermatogenesis, truncated kinases are still able to perform their function if the catalytic domain is maintained (Kierszenbaum, 2006). Similarly, the paternally-derived embryonic circRNAs predicted herein appear to be able to generate truncated forms of their canonical proteins. For instance, the theoretical protein products of the MTG1 and VRK1 derived circRNAs seem to correspond specifically to their catalytic regions. Therefore, since circRNAs are more stable and resistant, and possess longer half-lives than their linear cognates, could be feasible that sperm circRNAs would be maintained in their circular form to protect essential functions until fertilization and then be translated during the first steps of embryo development using the oocyte translational machinery.

In conclusion, our study has demonstrated that spermatozoa contain a complex population of circRNAs, including exonic and intronic circRNAs. During spermatogenesis, linear forms seem to be gradually degraded, while the circRNAs would be selectively retained in the human mature sperm and delivered at fertilization, holding potential functions for the developing embryo (Figure 7). Our results have highlighted that some zygote exonic circRNAs are paternally-derived, as well as that some SREs required to achieve a live birth and/or found altered in obese patients appear to be in fact circRNAs, supporting the assumption that sperm circRNAs are necessary to proper sperm functionality and harbor potential epigenetic roles beyond fertilization. The presence of intact ORFs in circRNAs reinforces the hypothesis that sperm circRNAs can be translated to functional peptides with a putative role in the embryogenesis. Further peptidomic studies targeting peptides derived from circRNAs, as well as functional studies would shed new light in elucidating their specific contribution to the developing embryo and possibly the offspring's health.

DATA AVAILABILITY

RNA-seq data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and datasets are available through GEO Series accession number GSE158778 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE158778).

SUPPLEMENTARY DATA

Supplementary Data has been deposited to OneDrive repository from the University of Barcelona. The data files can be downloaded at the following link until published: https://ubarcelona-my.sharepoint.com/:f:/r/personal/ada_soler_ub_edu/Documents/Treball%201?csf=1&web=1&e=d5QR4g

ACKNOWLEDGEMENTS

The authors are in depth thankful with Salzman Lab (Stanford University, Stanford, California, USA) for their teaching and advisement of circRNA prediction and validation methods. Human embryonic kidney (HEK) cells were provided by members of the neurophysiology group (Institute of Neurosciences, Department of Biomedical Sciences, Universitat de Barcelona, Spain). The authors recognize Raquel Ferreti and Alicia Diez for their assistance in the routine seminograms and sample collection.

FUNDING

This work was supported by the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad; Fondos FEDER 'Una manera de hacer Europa') [grant numbers PI13/00699 and PI16/00346 to R.O.]; Postdoctoral Fellowship from the Government of Catalonia (Generalitat de Catalunya, pla estratègic de recerca i innovació en salut, PERIS 2016-2020) [grant number SLT SLT002/16/00337 to M.J.]; PROLAB 2017 award from American Society for Biochemistry and Molecular Biology (ASBMB) to M.J.; and Sara Borrell Postdoctoral Fellowship from the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad, Acción Estratégica en Salud) [grant number CD17/00109 to J.C.].

CONFLICT OF INTEREST

The authors declare no competing financial interest.

REFERENCES

Aktaş T, Avşar Ilık İ, Maticzka D, Bhardwaj V, Pessoa Rodrigues C, Mittler G, Manke T, Backofen R, Akhtar A. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. Nature 2017;544:115–119.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene Ontology: tool for the unification of biology. Nat Genet 2000;25:25–29.

Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S. circRNA Biogenesis Competes with Pre-mRNA Splicing. Mol Cell 2014;56:55–66.

Aufiero S, Hoogenhof MMG van den, Reckman YJ, Beqqali A, Made I van der, Kluin J, Khan MAF, Pinto YM, Creemers EE. Cardiac circRNAs arise mainly from constitutive exons rather than alternatively spliced exons. RNA 2018;24:815–827.

Bachmayr-Heyda A, Reiner AT, Auer K, Sukhbaatar N, Aust S, Bachleitner-Hofmann T, Mesteri I, Grunt TW, Zeillinger R, Pils D. Correlation of circular RNA abundance with proliferation – exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis and normal human tissues. Sci Rep 2015;5:8057.

Barrett SP, Salzman J. Circular RNAs: analysis, expression and potential functions. Development 2016;143:1838–1847.

Barrett SP, Wang PL, Salzman J. Circular RNA biogenesis can proceed through an exon-containing lariat precursor. Elife 2015;4:1–18.

Castillo J, Jodar M, Oliva R. The contribution of human sperm proteins to the development and epigenome of the preimplantation embryo. Hum Reprod Update 2018;24:535–555. England.

Chen Q, Yan M, Cao Z, Li X, Zhang Y, Shi J, Feng GH, Peng H, Zhang X, Zhang Y, et al. Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. Science 2016;351:397–400.

Chioccarelli T, Manfrevola F, Ferraro B, Sellitto C, Cobellis G, Migliaccio M, Fasano S, Pierantoni R, Chianese R. Expression patterns of circular RNAs in high quality and poor quality human spermatozoa. Front Endocrinol (Lausanne) 2019;10:435. Frontiers Media S.A.

Chioccarelli T, Pierantoni R, Manfrevola F, Porreca V, Fasano S, Chianese R, Cobellis G. Histone Post-Translational Modifications and CircRNAs in Mouse and Human Spermatozoa: Potential Epigenetic Marks to Assess Human Sperm Quality. J Clin Med 2020;9:640. Multidisciplinary Digital Publishing Institute.

Choi YH, Lim JK, Jeong MW, Kim KT. HnRNP A1 phosphorylated by VRK1 stimulates telomerase and its binding to telomeric DNA sequence. Nucleic Acids Res 2012;40:8499–8518. Oxford University Press.

Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, Roslan S, Schreiber AW, Gregory PA, Goodall GJ. The RNA Binding Protein Quaking Regulates Formation of circRNAs. Cell 2015;160:1125–1134.

Dang Y, Yan L, Hu B, Fan X, Ren Y, Li R, Lian Y, Yan J, Li Q, Zhang Y, et al. Tracing the expression of circular RNAs in human pre-implantation embryos. Genome Biol 2016;17:130.

Development Core Team. R: A Language and Environment for Statistical Computing. R Found Stat Comput 2018;

Dong W-W, Li H-M, Qing X-R, Huang D-H, Li H-G. Identification and characterization of human testis derived circular RNAs and their existence in seminal plasma. Sci Rep 2016;6:39080. Nature Publishing Group.

Donkin I, Versteyhe S, Ingerslev LR, Qian K, Mechta M, Nordkap L, Mortensen B, Appel EVR, Jørgensen N, Kristiansen VB, et al. Obesity and bariatric surgery drive epigenetic variation of spermatozoa in humans. Cell Metab 2016;23:369–378.

Dori M, Bicciato S. Integration of bioinformatic predictions and experimental data to identify circRNA-miRNA associations. Genes (Basel) 2019;10:642. MDPI AG.

Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–210. Oxford University Press.

Enuka Y, Lauriola M, Feldman ME, Sas-Chen A, Ulitsky I, Yarden Y. Circular RNAs are long-lived and display only minimal early alterations in response to a growth factor. Nucleic Acids Res 2016;44:1370–1383.

Errichelli L, Dini Modigliani S, Laneve P, Colantoni A, Legnini I, Capauto D, Rosa A, Santis R De, Scarfò R, Peruzzi G, et al. FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. Nat Commun 2017;8:14741.

Gòdia M, Castelló A, Rocco M, Cabrera B, Rodríguez-Gil JE, Balasch S, Lewis C, Sánchez A, Clop A. Identification of circular RNAs in porcine sperm and evaluation of their relation to sperm motility. Sci Rep 2020;10:1–11. Nature Research.

Gòdia M, Swanson G, Krawetz SA. A history of why fathers' RNA matters†. Biol Reprod 2018;99:147–159. Narnia.

Goodrich RJ, Anton E, Krawetz SA. Isolating mRNA and small noncoding RNAs from human sperm. Methods Mol Biol 2013;927:385–396.

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 2010;38:W695–W699. Oxford Academic.

Greene J, Baird AM, Brady L, Lim M, Gray SG, McDermott R, Finn SP. Circular RNAs: Biogenesis, function and role in human diseases. Front Mol Biosci 2017;4:38.

Gruner H, Cortés-López M, Cooper DA, Bauer M, Miura P. CircRNA accumulation in the aging mouse brain. Sci Rep 2016;6:38907.

Guo JU, Agarwal V, Guo H, Bartel DP. Expanded identification and characterization of mammalian circular RNAs. Genome Biol 2014;15:409. BioMed Central Ltd.
Haddad G, Lorenzen JM. Biogenesis and Function of Circular RNAs in Health and in Disease. Front Pharmacol 2019;10:428.
Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. Natural RNA circles function as efficient microRNA sponges. Nature 2013;495:384–388.
Hansen TB, Wiklund ED, Bramsen JB, Villadsen SB, Statham AL, Clark SJ, Kjems J. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. EMBO J 2011;30:4414–4422.
Huang C, Liang D, Tatomer DC, Wilusz JE. A length-dependent evolutionarily conserved pathway controls nuclear export of circular RNAs. Genes Dev 2018;32:639–644.
Huang J, Zhou Q, Li Y. Circular RNAs in gynecological disease: Promising biomarkers and diagnostic targets. Biosci Rep 2019;39:e1592. Portland Press Ltd.
Iparraguirre L, Prada-Luengo I, Regenberg B, Otaegui D. To Be or Not to Be: Circular RNAs or mRNAs From Circular DNAs? Front Genet 2019;10:940. Frontiers Media S.A.
Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR, Piechotta M, Levanon EY, Landthaler M, Dieterich C, et al. Analysis of Intron Sequences Reveals Hallmarks of Circular RNA Biogenesis in Animals. Cell Rep 2015;10:170–177.
Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE. Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA 2013;19:141–157.
Jodar M. Sperm and seminal plasma RNAs: What roles do they play beyond fertilization? Reproduction 2019;158:R113–R123.
Jodar M, Attardo-Parrinello C, Soler-Ventura A, Barrachina F, Delgado-Dueñas D, Cívico S, Calafell JM, Ballescà JL, Oliva R. Sperm proteomic changes associated with early embryo quality after ICSI. Reprod Biomed Online 2020;40:698–708. Elsevier Ltd.
Jodar M, Selvaraju S, Sendler E, Diamond MP, Krawetz SA. The presence, role and clinical use of spermatozoal RNAs. Hum Reprod Update 2013;19:604–624.
Jodar M, Sendler E, Moskovtsev SI, Librach CL, Goodrich R, Swanson S, Hauser R, Diamond MP, Krawetz SA. Absence of sperm RNA elements correlates with idiopathic male infertility. Sci Transl Med 2015;7:295re6.
Kierszenbaum AL. Tyrosine protein kinases and spermatogenesis: Truncation matters. Mol Reprod Dev 2006;73:399–403. Mol Reprod Dev.
Kleaveland B, Shi CY, Stefano J, Bartel DP. A Network of Noncoding Regulatory RNAs Acts in the Mammalian Brain. Cell 2018;174:350-362.e17.
Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. Bioinformatics 2007;23:1289–1291. Bioinformatics.

Kramer MC, Liang D, Tatomer DC, Gold B, March ZM, Cherry S, Wilusz JE. Combinatorial control of Drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins. Genes Dev 2015;29:2168–2182.

Krawetz SA, Kruger A, Lalancette C, Tagett R, Anton E, Draghici S, Diamond MP. A survey of small RNAs in human sperm. Hum Reprod 2011;26:3401–3412.

Kristensen LS, Andersen MS, Stagsted LVW, Ebbesen KK, Hansen TB, Kjems J. The biogenesis, biology and characterization of circular RNAs. Nat Rev Genet 2019;20:675–691.

Lau JW, Lehnert E, Sethi A, Malhotra R, Kaushik G, Onder Z, Groves-Kirkby N, Mihajlovic A, DiGiovanna J, Srdic M, et al. The Cancer Genomics Cloud: Collaborative, Reproducible, and Democratized—A New Paradigm in Large-Scale Computational Research. Cancer Res 2017;77:e3–e6.

Legnini I, Timoteo G Di, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M, et al. Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. Mol Cell 2017;66:22-37.e9.

Li X, Liu CX, Xue W, Zhang Y, Jiang S, Yin QF, Wei J, Yao RW, Yang L, Chen LL. Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection. Mol Cell 2017;67:214-227.e7.

Li X, Zhu L, Yang A, Lin J, Tang F, Jin S, Wei Z, Li J, Jin Y. Calcineurin-NFAT Signaling Critically Regulates Early Lineage Specification in Mouse Embryonic Stem Cells and Embryos. Cell Stem Cell 2011;8:46–58.

Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L, et al. Exon-intron circular RNAs regulate transcription in the nucleus. Nat Struct Mol Biol 2015;22:256–264.

Liang D, Wilusz JE. Short intronic repeat sequences facilitate circular RNA production. Genes Dev 2014;28:2233–2247.

Medvedeva YA, Lennartsson A, Ehsani R, Kulakovskiy I V., Vorontsov IE, Panahandeh P, Khimulya G, Kasukawa T, Drabløs F. EpiFactors: A comprehensive database of human epigenetic factors and complexes. Database 2015;7:bav067. Narnia.

Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature 2013;495:333–338.

Ostermeier GC, Miller D, Huntriss JD, Diamond MP, Krawetz SA. Delivering spermatozoan RNA to the oocyte. Nature 2004;429:154–154.

Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E, et al. Translation of CircRNAs. Mol Cell 2017;66:9-21.e7.

Panda AC, De S, Grammatikakis I, Munk R, Yang X, Piao Y, Dudekula DB, Abdelmohsen K, Gorospe M. High-purity circular RNA isolation method (RPAD) reveals vast collection of intronic circRNAs. Nucleic Acids Res 2017;45:e116–e116.

Park OH, Ha H, Lee Y, Boo SH, Kwon DH, Song HK, Kim YK. Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex. Mol Cell 2019;74:494-507.e8.

Patop IL, Wüst S, Kadener S. Past, present, and future of circ RNA s. EMBO J 2019;38:e100836. John Wiley & Sons, Ltd.

Ragusa M, Barbagallo D, Chioccarelli T, Manfrevola F, Cobellis G, Pietro C Di, Brex D, Battaglia R, Fasano S, Ferraro B, et al. CircNAPEPLD is expressed in human and murine spermatozoa and physically interacts with oocyte miRNAs. RNA Biol 2019;16:1237–1248. Taylor and Francis Inc.

Rybak-Wolf A, Stottmeister C, Glažar P, Jens M, Pino N, Giusti S, Hanan M, Behm M, Bartok O, Ashwal-Fluss R, et al. Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. Mol Cell 2015;58:870–885.

Salzman J, Chen RE, Olsen MN, Wang PL, Brown PO. Cell-Type Specific Features of Circular RNA Expression. In Moran J V., editor. PLoS Genet 2013;9:e1003777.

Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. In Preiss T, editor. PLoS One 2012;7:e30733.

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T. The RIN: An RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol 2006;7:3.

Sendler E, Johnson GD, Mao S, Goodrich RJ, Diamond MP, Hauser R, Krawetz SA. Stability, delivery and functions of human sperm RNAs at fertilization. Nucleic Acids Res 2013;41:4104–4117.

Seytanoglu A, Alsomali NI, Valori CF, McGown A, Kim HR, Ning K, Ramesh T, Sharrack B, Wood JD, Azzouz M. Deficiency in the mRNA export mediator Gle1 impairs Schwann cell development in the zebrafish embryo. Neuroscience 2016;322:287–297.

Sharma U, Conine CC, Shea JM, Boskovic A, Derr AG, Bing XY, Belleannee C, Kucukural A, Serra RW, Sun F, et al. Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. Science 2016;351:391–396.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 2011;7:539. John Wiley & Sons, Ltd.

Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung L-H, Bindereif A. Exon Circularization Requires Canonical Splice Signals. Cell Rep 2015;10:103–111.

Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Iny Stein T, Nudel R, Lieder I, Mazor Y, et al. The GeneCards suite: From gene data mining to disease genome sequence analyses. Curr Protoc Bioinforma 2016;20:1.30.1-1.30.33. John Wiley and Sons Inc.

Sun X, Wang L, Ding J, Wang Y, Wang J, Zhang X, Che Y, Liu Z, Zhang X, Ye J, et al. Integrative analysis of Arabidopsis thaliana transcriptomics reveals intuitive splicing mechanism for circular RNA. FEBS Lett 2016;590:3510–3516.

Swanson GM, Estill M, Diamond MP, Legro RS, Coutifaris C, Barnhart KT, Huang H, Hansen KR, Trussell JC, Coward RM, et al. Human chromatin remodeler cofactor, RNA interactor, eraser and writer sperm RNAs responding to obesity. Epigenetics 2020;15:32–46.

Szabo L, Morey R, Palpant NJ, Wang PL, Afari N, Jiang C, Parast MM, Murry CE, Laurent LC, Salzman J. Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. Genome Biol 2015;16:126.

Szabo L, Salzman J. Detecting circular RNAs: bioinformatic and experimental challenges. Nat Rev Genet 2016;17:679–692.

Tang C, Xie Y, Yu T, Liu N, Wang Z, Woolsey RJ, Tang Y, Zhang X, Qin W, Zhang Y, et al. m6A-dependent biogenesis of circular RNAs in male germ cells. Cell Res 2020;30:211–228. Springer Nature.

The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res 2019;47:D330–D338.

Turner S, Wong HP, Rai J, Hartshorne GM. Telomere lengths in human oocytes, cleavage stage embryos and blastocysts. Mol Hum Reprod 2010;16:685–694.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. Primer3-new capabilities and interfaces. Nucleic Acids Res 2012;40:e115–e115. Oxford Academic.

Venø MT, Hansen TB, Venø ST, Clausen BH, Grebing M, Finsen B, Holm IE, Kjems J. Spatio-temporal regulation of circular RNA expression during porcine embryonic brain development. Genome Biol 2015;16:245.

Wang Y, Wang Z. Efficient backsplicing produces translatable circular mRNAs. RNA 2015;21:172–179.

Westholm JO, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P, Celniker SE, Graveley BR, Lai EC. Genome-wide Analysis of Drosophila Circular RNAs Reveals Their Structural and Sequence Properties and Age-Dependent Neural Accumulation. Cell Rep 2014;9:1966–1980.

Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Tatusova TA, et al. Database resources of the national center for biotechnology. Nucleic Acids Res 2003;31:28–33. Oxford University Press.

White K, Yang P, Li L, Farshori A, Medina AE, Zielke HR. Effect of Postmortem Interval and Years in Storage on RNA Quality of Tissue at a Repository of the NIH NeuroBioBank. Biopreserv Biobank 2018;16:148–157.

World Health Organization. Laboratory manual for the examination and processing of human semen. Cambridge Cambridge Univ Press 2010; World Health Organization.

Yan N, Xu H, Zhang J, Xu L, Zhang Y, Zhang L, Xu Y, Zhang F. Circular RNA profile indicates circular RNA VRK1 is negatively related with breast cancer stem cells. Oncotarget 2017;8:95704–95718.

Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, Jin Y, Yang Y, Chen LL, Wang Y, et al. Extensive translation of circular RNAs driven by N 6 -methyladenosine. Cell Res 2017;27:626–641. Nature Publishing Group.

Yu CY, Li TC, Wu YY, Yeh CH, Chiang W, Chuang CY, Kuo HC. The circular RNA circBIRC6 participates in the molecular circuitry controlling human pluripotency. Nat Commun 2017;8:1149.

Zeng X, Lin W, Guo M, Zou Q. A comprehensive overview and evaluation of circular RNA detection tools. PLoS Comput Biol 2017;13:e1005420. Public Library of Science.

Zhang X, Wang W, Zhu W, Dong J, Cheng Y, Yin Z, Shen F. Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. Int J Mol Sci 2019;20:5573. MDPI AG.

Zhang XO, Wang H Bin, Zhang Y, Lu X, Chen LL, Yang L. Complementary Sequence-Mediated Exon Circularization. Cell 2014;159:134–147.

Zhang Y, Zhang X-O, Chen T, Xiang J-F, Yin Q-F, Xing Y-H, Zhu S, Yang L, Chen L-L. Circular Intronic Long Noncoding RNAs. Mol Cell 2013;51:792–806.

3.2. Treball 2

Anàlisi i extracció de les protamines de l'espermatozoide de mamífers: Protocol detallat pas a pas i breu revisió de les alteracions de les protamines

Protein and Peptide Letters

2018;25(5):424-433

PMID: 29651936

Mammalian Sperm Protamine Extraction and Analysis: A Step-By-Step Detailed Protocol and Brief Review of Protamine Alterations.

Soler-Ventura A*, Castillo J*, de la Iglesia A, Jodar M, Barrachina F, Ballescà JL, Oliva R.

* Aquests autors han contribuït per igual a aquest treball

Anàlisi i extracció de les protamines de l'espermatozoide de mamífers: Protocol detallat pas a pas i breu revisió de les alteracions de les protamines

Objectiu: Detallar de forma extensa el protocol d'extracció de protamines de l'espermatozoide de mamífers pas a pas, optimitzar-lo per a l'anàlisi posterior de protamines mitjançant cromatografia líquida seguit d'espectrometria de masses en tàndem (LC-MS/MS) i realitzar una breu revisió de l'associació de les alteracions de les protamines amb la infertilitat masculina.

Resultats: En aquesta revisió es detallen els diferents passos per l'extracció de protamines de l'espermatozoide. Breument, primerament hi ha un rentat dels espermatozoides, seguit de la permeabilització de la seva membrana i d'un xoc osmòtic. A continuació es descondensa i precipita la cromatina. Seguidament, s'extrauen les proteïnes bàsiques del DNA. El DNA aïllat es pot hidrolitzar i quantificar i, les proteïnes bàsiques es precipiten i mantenen en condicions reductives per la seva posterior anàlisi. Finalment, es fa una separació de les protamines mitjançant electroforesi en gel de poliacrilamida en condicions àcides i en presència d'urea, seguit de la seva tinció, el que permet visualitzar i quantificar la protamina 1 (P1) i els membres de la família de protamina 2 (P2) en comparació a un estàndard de protamines. En el cas que la finalitat de l'aïllament de les protamines sigui la seva posterior anàlisi per LC-MS/MS, després de la permeabilització de la membrana s'eliminen les histones i altres proteïnes bàsiques amb tres rentats d'àcid clorhídric (HCl) i s'utilitzen reactius ultrapurs i compatibles amb espectrometria de masses (com ara l'àcid tricloroacètic (TCA) i l'acetona de qualitat especial), així com condicions apropiades (tals com tubs de baixa retenció de proteïnes). Una petita part de l'extracció s'utilitzarà per visualitzar i quantificar mitjançant una electroforesi en gel àcid – urea i la resta es guardarà liofilitzada fins a la seva anàlisi. En aquest treball, també, s'han resumit els estudis, que mitjançant l'electroforesi en gel àcid – urea, eina convencional d'anàlisi de protamines, connecten les alteracions de les protamines amb la infertilitat masculina. El resultat de la revisió de l'associació de les alteracions de les protamines amb la infertilitat masculina posa de manifest nombrosos treballs descrivint alteracions en el ràtio P1/P2 i pre-P2/P2 correlacionades a un augment en el dany al DNA de l'espermatozoide. També, alteracions en aquests ràtios s'han vinculat al desenvolupament embrionari, com ara una baixa taxa de fertilització, baixa taxa d'implantació, baixa taxa de resultat d'embaràs i baixa qualitat embrionària. De manera similar, s'han correlacionat alteracions en el ràtio P1/P2 i pre-P2/P2 a alteracions en els paràmetres seminals.

Conclusions: S'ha descrit detalladament el protocol d'extracció de protamines de l'espermatozoide de mamífers amb múltiples modificacions. Addicionalment, s'ha optimitzat el protocol per a l'anàlisi de protamines per LC-MS/MS. A més, també s'ha fet una breu revisió dels diferents estudis publicats fins a l'actualitat sobre alteracions de les protamines i la infertilitat masculina.

Send Orders for Print-Reprints and e-prints to reprints@benthamscience.ae

Protein & Peptide Letters, **2018**, *25*, 424-433

RESEARCH ARTICLE

Mammalian Sperm Protamine Extraction and Analysis: A Step-By-Step Detailed Protocol and Brief Review of Protamine Alterations

Ada Soler-Ventura[a,#], Judit Castillo[a,#], Alberto de la Iglesia[a], Meritxell Jodar[a], Ferran Barrachina[a], Josep Lluís Ballescà[b] and Rafael Oliva[a,*]

[a]*Molecular Biology of Reproduction and Development Research Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain, and Biochemistry and Molecular Genetics Service, Hospital Clínic, Barcelona, Spain;* [b]*Clinic Institute of Gynecology, Obstetrics and Neonatology, Hospital Clínic, Barcelona, Spain*

ARTICLE HISTORY

Received: July 19, 2017
Revised: September 21, 2017
Accepted: January 5, 2018

DOI:
10.2174/0929866525666180412155205

Abstract: ***Background*:** Protamines are the most abundant sperm nuclear proteins and pack approximately the 92-98% of the mammalian sperm DNA. In mammals, two types of protamines have been described, the protamine 1 (P1) and the protamine 2 (P2) family. The deregulation of the relative P1/P2 ratio has been correlated to DNA damage, alterations in seminal parameters, and low success rate of assisted reproduction techniques. Additionally, the extraction and analysis of protamines have been important to understand the fundamental aspects of the sperm chromatin structure and function, protamine sequence conservation among species, and sperm chromatin alterations present in infertile males. However, protamines show a particular chemical nature due to its special amino acid sequence, extremely rich in arginine and cysteine residues. Because of these peculiar characteristics of protamines, their extraction and analysis is not as straightforward as the analysis of other chromatin-associated proteins, for which many detailed protocols are already available.

***Conclusion*:** A step-by-step protocol was needed to facilitate protamine analysis to researchers interested in their implementation. Therefore, in order to contribute to fulfill this need, here we provide a detailed protocol, which should be useful to research teams and laboratories interested in the protamine field. In addition, we also briefly review the different studies published so far on protamine alterations and male infertility.

Keywords: Protamines, sperm, protein isolation, male infertility, acid-urea electrophoresis, nuclear proteins, sperm chromatin.

1. INTRODUCTION

Protamines are the most abundant sperm nuclear proteins associated to the chromatin in many species. These are small and extremely basic proteins that package the major part of the paternal genome, although in a different proportion depending on the species. For instance, protamines condense approximately the 92% of the human sperm DNA (Figure **1A**), and about the 98% in mouse [1, 2]. Protamines show particular physical-chemical properties due to their special amino acid sequence, which is extremely rich in positively-charged arginine residues (50-70%), allowing the formation of highly-condensed toroidal complexes together with the negatively-charged paternal DNA. Additionally, mammalian protamines are rich in cysteine residues, which promotes the generation of disulfide bonds and zinc bridges among intra- and inter-protamine molecules, resulting in a compact toroidal nucleoprotamine complex (Figure **1A**) [2, 3].

Protamines are bound and complexed with the sperm genome in the latest phase of spermatogenesis called spermiogenesis, during the nucleohistone-to-nucleoprotamine transition (Figure **1A**). This is a highly epigenetically regulated process in which histones are firstly substituted by histone variants, possibly contributing to an appropriate and more accessible state of chromatin compaction necessary for the incorporation of transition proteins (TNPs) [4-9]. Recent evidence suggests that, in contrast to what was previously supposed, TNPs would not directly replace histones, but rather would be involved in the recruitment and processing of protamines, being protamines the potential true leaders of the histone replacement [10-12]. The histone exchange process is required for the sperm chromatin to become highly compacted, protecting the paternal DNA from nucleases, and contributing to the acquisition of the sperm hydrodynamic shape needed for a proper functionality [1, 5, 9, 11-13].

Two types of protamines have been described in mammals: the protamine 1 (P1) and the protamine 2 (P2) family. Whereas P1 is synthesized as a mature form, the P2 family is

*Address correspondence to this author at the Molecular Biology of Reproduction and Development Research Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain, and Biochemistry and Molecular Genetics Service, Hospital Clínic, Barcelona, Spain; Tel/Fax: +00340-934-02-1877; E-mail: roliva@ub.edu
[#]*A.S-V and J.C contributed equally to this work.*

0929-8665/18 $58.00+.00

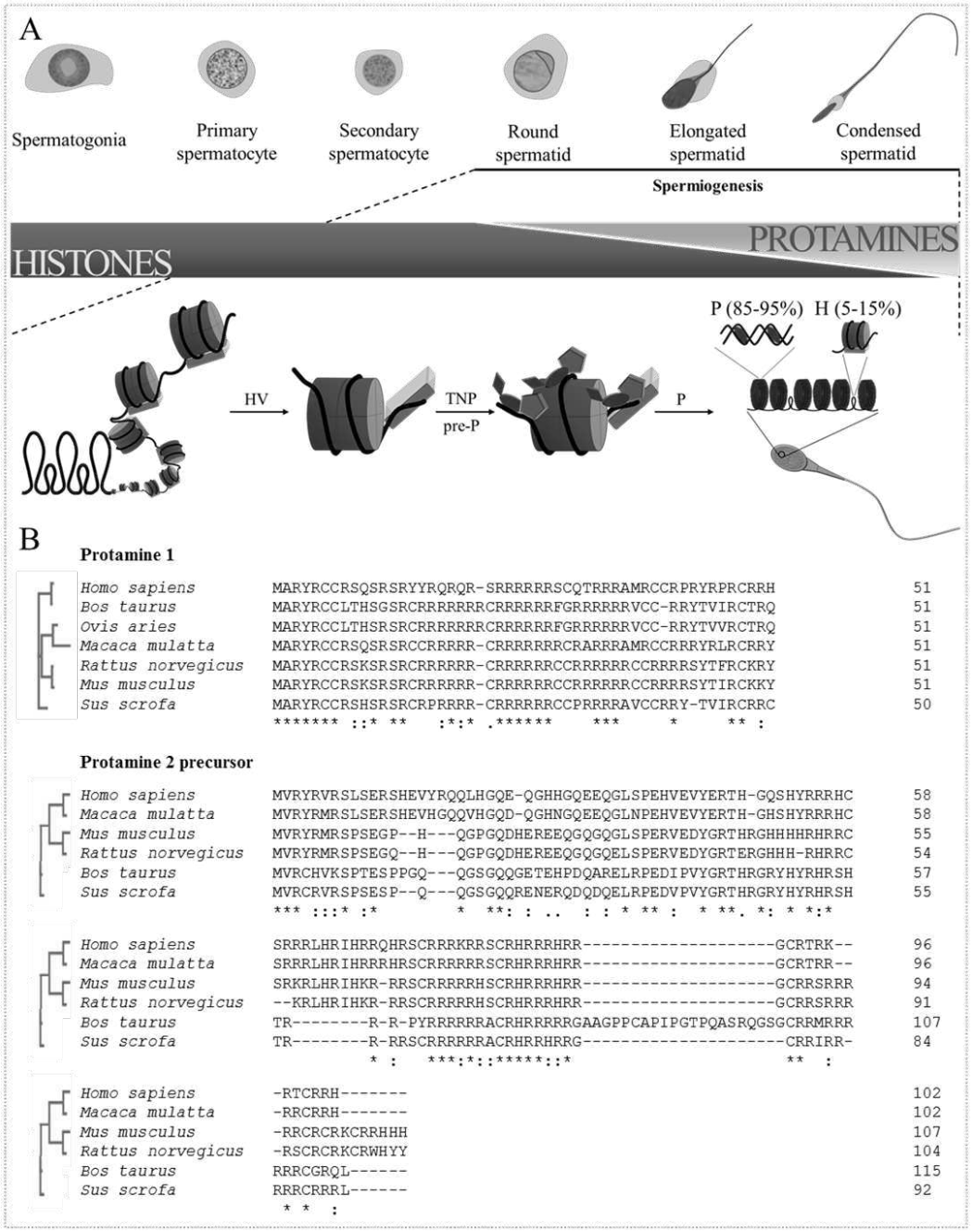


Figure 1. A) Nucleohistone-to-nucleoprotamine transition occurring during human spermiogenesis. H: Histones, HV: Histone Variants, TNP: Transition Nuclear Proteins, pre-P: Protamine precursor, P: protamines. Model not drawn at scale. **B**) "CLUSTAL O" (http://www.ebi.ac.uk/Tools/msa/clustalo/) of Protamine 1 (upper part) and Protamine 2 precursor (lower part) for different mammal species. * single, fully conserved residue, : conservation of strongly similar properties, . conservation of weakly similar properties.

generated from the proteolysis of the protamine 2 precursor (pre-P2), resulting in the three different components of the P2 family (P2, P3, and P4). P2 family components differ only by 1-4 amino acid residues on the N-terminal extension, being component P2 the most abundant [5, 14]. All mammal species harbor P1 in spermatozoa; however, the P2 family is solely expressed by some mammal species, such as humans and mice (Figure **1B**) [5, 15]. In fact, protamines show one of the highest rates of evolutionary variation, most probably caused by a positive Darwinian selection [1, 16-19], which results in substantial amino acid sequence variation among closely related species (Figure **1B**). Although the overall number of arginine residues is maintained around 50% among mammals, the total number of protamine amino acids and the particular positions in the sequence changes considerably (Figure **1B**) [20].

The negatively charged DNA and the coupled positively charged protamines were the first components described in the sperm chromatin. The first evidence of the existence of protamines was provided by Friedreich Miescher in 1874 [21]. Miescher isolated a basic component from salmon sperm that he called protamine, and observed that it was coupled to what he first called nuclein but that later became known as DNA [22]. However, the principles for protamine extraction and analysis were developed much later, with the goal to obtain large quantities of protamines, enough to determine their amino acid sequences and modifications [1, 6, 15, 23-29]. However, these initial reported methods were quite time-consuming, requiring several days to complete the extraction and purification and, while they were suitable to determine the amino acid sequence, these procedures were inefficient for the fast and routine analysis of clinical samples.

A systematic analysis was then undertaken by our group to determine the impact of reducing the different incubation times involved in each protocol step [30]. This systematic analysis allowed us to optimize the overall time required, resulting in a few-hours procedure that opened up the possibility to start analyzing more efficiently large numbers of sperm clinical samples from infertile patients [31-48]. More recently, the advent of high-throughput proteomic methodologies based on mass spectrometry (MS) has also enabled to detect protamines and analyze their potential post-translational modifications [49, 50].

The extraction and analysis of protamines are of interest to understand the fundamental aspects of the sperm chromatin structure and function [1, 11, 12, 15, 51-53], protamine evolution [16-20, 54-60], and sperm chromatin alterations in infertile males [30, 31, 34, 42, 45, 61-64]. Also protamines are used in pharmaceutics as a heparin antagonist and more recently in DNA/RNA nanoparticle complexes for gene transfer to cells [65-67]. However, due to the particular chemical nature of protamines, their extraction and analysis is not as straightforward as the analysis of histones and other chromatin associated proteins from somatic cells, for which many detailed protocols are already available [68-70]. Additionally, in the case of protamines, the methods published so far for their extraction, purification and analysis are generally quite succinct [30, 31, 71, 72]. From a practical perspective, the technical procedure for protamine extraction requires some critical steps, which deserves a detailed step-by-step description. Therefore, here we provide a technical review including the chemical principles involved in the extraction and analysis of protamines, the methods available, and a very comprehensive protocol to allow the successful implementation of protamine extraction and analysis. This article is focused on the direct quantification of protamines at the protein level in mature sperm, and does not intent to cover the protamine quantification of mRNA, for which the reader is referred to other articles or reviews [14, 47], or indirect chromatin staining procedures [35, 73, 74]. A revision of the studies published so far about protamine alterations and male infertility, as well as the detailed protocol for the extraction and analysis of protamines, including explanations about the critical steps and protocol variants depending on the research goal, are described herein.

2. PROTAMINES AND MALE INFERTILITY

The relative ratio of the abundance of P1 and P2 family as a measure for the evaluation of the sperm chromatin maturity and the normality/abnormality of the sperm has been widely studied. For fertile males, it was proposed a P1/P2 ratio around 1 (range of 0.8-1.2) [51, 61]. However, prospective studies in the general population suggested the presence of a broader range for P1/P2 ratio that can oscillate between 0.5 and 1.5 [31, 34, 43]. A deregulation in the P1/P2 ratio has been correlated to DNA damage, seminal parameters alterations and a low success rate of assisted reproduction techniques (Table **1**) [31, 33, 36-39, 41, 42, 44-46, 62, 63, 75-80]. In addition, a recent meta-analysis from Ni *et al.* (2016), comprising data from nine different studies and comparing fertile and infertile patients, concluded that a significantly increase of P1/P2 ratio is related to subfertility [47].

Although the recent meta-analysis performed by Ni *et al.* (2016) could not establish a relationship between altered P1/P2 ratio and DNA damage [47], other studies have demonstrated that alterations on the P1/P2 ratio and reductions on the pre-P2/P2 ratio are linked to an increase in DNA damage and reactive oxygen species. These studies suggest that sperm DNA becomes more accessible to nuclease activity due to the altered P1/P2 ratio, causing the damage of the DNA (Table **1**) [37, 38, 41, 44-46, 75, 76]. On the other hand, it has been shown that P2 is deregulated more frequently than P1, suggesting that a P2 deregulation is usually responsible for the P1/P2 alterations. Low sperm counts and sperm motility, and/or abnormal sperm morphology have been also correlated with abnormal P1/P2 ratios (Table **1**) [34, 36, 37, 42, 62, 63, 78-80]. Similarly, some studies have demonstrated that the incomplete processing of the pre-P2, reflected by a decreased pre-P2/P2 ratio, or the total absence of P2 are also related to altered seminal parameters (Table **1**) [31, 33, 38, 42].

Moreover, several studies have shown correlations between an altered P1/P2 ratio and low fertilization or low implantation rates, low embryo quality scores, and low pregnancy outcomes, using assisted reproduction techniques such as *in vitro* fertilization (IVF) or intracytoplasmic sperm injection (ICSI) (Table **1**) [33, 36, 39, 42, 46, 77, 78]. Similarly, the absence of P2 has been linked to a low sperm penetration ability [33], and a decreased pre-P2/P2 ratio has been correlated to low implantation rates and low pregnancy outcomes [42]. Overall, these studies suggest that protamine deregulation could interfere in fertilization and early embryo development.

3. MATERIALS

3.1. Reagents

- 2-Amino-2-(hydroxymethyl)-1,3-propanediol hydrochloric acid (Trizma® base HCl) (Sigma-Aldrich)
- 2-propanol (Merck Millipore)
- Acetic Acid glacial (Panreac)
- Acetone (Merck Millipore)

Table 1. Altered P1/P2 ratio in subfertile patients. Publications showing correlations of protamine P1/P2 ratio with DNA damage, assisted reproduction techniques outcome and seminal parameters.

Study	P1/P2 Ratio	Outcome
Correlation with DNA Damage		
Ribas-Maynou *et al.* (2015) [75]	Increased P1/P2	Increased DNA damage
García-Peiró *et al.* (2011) [44]	Increased P1/P2	Increased DNA damage
Castillo *et al.* (2011) [45]	Decreased P1/P2	Increased DNA damage
Simon *et al.* (2011) [46]	Increased P1/P2	Increased DNA damage
Hammadeh *et al.* (2010) [41]	Increased P1/P2	Increased reactive oxygen species
Aoki *et al.* (2006) [76]	Altered P1/P2	Increased DNA damage
Torregrosa *et al.* (2006) [38]	Decreased pre-P2/P2	Increased DNA damage
Aoki *et al.* (2005) [37]	Decreased P1/P2	Increased DNA damage
Correlation with Assisted Reproduction Techniques		
Simon *et al.* (2011) [46]	Decreased P1/P2	Low fertilization rate (IVF)
de Mateo *et al.* (2009) [42]	Decreased P1/P2	Low fertilization rate (IVF); Low implantation rate (IVF and/or ICSI); Low pregnancy outcome (IVF and/or ICSI)
	Decreased pre-P2/P2	Low implantation rate (IVF and/or ICSI); Low pregnancy outcome (IVF and/or ICSI)
Aoki *et al.* (2006) [39]	Altered P1/P2	Low fertilization rate (IVF)
	Decreased P1/P2	Low chemical-pregnancy and clinical-pregnancy rates (IVF and/or ICSI)
Aoki *et al.* (2005) [36]	Decreased P1/P2	Low fertilization rate (IVF and ICSI)
Nasr-Esfahani *et al.* (2004) [77]	Increased P1/P2	Low fertilization rate (ICSI); Low embryo quality score in day 3 (ICSI)
Khara *et al.* (1997) [78]	Altered P1/P2	Low fertilization rate (IVF)
Carrell and Liu (2001) [33]	No P2	Low sperm penetration ability (IVF)
Correlation with Seminal Parameters		
Simon *et al.* (2014) [79]	Altered P1/P2	Lower sperm count; Lower semen volume
Hamad *et al.* (2014) [80]	Increased P1/P2	Lower sperm count; Lower sperm vitality
de Mateo *et al.* (2009) [42]	Decreased P1/P2	Lower sperm motility
	Decreased pre-P2/P2	Lower sperm count; Lower sperm motility
Torregrosa *et al.* (2006) [38]	Decreased pre-P2/P2	Lower sperm count; Lower sperm motility; Abnormal sperm morphology
Aoki *et al.* (2005) [37]	Altered P1/P2	Lower sperm count; Lower sperm motility; Abnormal sperm morphology
Mengual *et al.* (2003) [34]	Increased P1/P2	Lower sperm count
Khara *et al.* (1997) [78]	Altered P1/P2	Lower sperm count; Lower sperm motility; Abnormal sperm morphology
de Yebra *et al.* (1993) [31]	Altered P1/P2	Lower sperm count
	No P2	Lower sperm count; Lower sperm motility; Abnormal sperm morphology
Aoki *et al.* (2005) [36]	Altered P1/P2	Lower sperm count; Lower sperm motility; Abnormal sperm morphology
	Decreased P1/P2	Abnormal sperm head morphology
Carrell and Liu (2001) [33]	No P2	Lower sperm motility; Abnormal sperm morphology
Bach *et al.* (1990) [63]	Altered P1/P2	Altered seminal parameters
Lescoat *et al.* (1988) [62]	Altered P1/P2	Altered seminal parameters

P1: Protamine 1; P2: Protamine 2; pre-P2: Protamine 2 Precursor; IVF: *In vitro* Fertilization; ICSI: Intracytoplasmic Sperm Injection.

- Acrylamide (Sigma-Aldrich)
- Ammonium Persulfate (APS) (Serva)
- DL-Dithiothreitol (DTT) (Sigma-Aldrich)
- Ethanol Absolute (Merck Millipore)
- Ethylenedinitrilotetraacetic acid, disodium salt dihydrate (EDTA) (Merck Millipore)
- *EZBlue*™ gel staining reagent (Sigma-Aldrich)
- Guanidinium chloride (Merck Millipore)
- HAM's F10 1x (Gibco by life technologies)
- Hydrochloric Acid (HCl) fuming 37% (Merck Millipore)
- Magnesium chloride hexahydrate ($MgCl_2$) (Merck Millipore)
- Methanol (Panreac)
- N,N′-Methylene bis-acrylamide (Sigma-Aldrich)
- Perchloric acid (PCA) (Merck Millipore)
- Phenylmethanesulfonyl fluoride (PMSF) (Sigma-Aldrich)
- Phosphate buffered saline (PBS) (Sigma-Aldrich)
- Pure Sperm® Buffer (Nidacon)
- Tetramethylethylenediamine (TEMED) (GE-Healthcare Life Sciences)
- Trichloroacetic acid (TCA) (Merck Millipore)
- Triton X-100 (Sigma-Aldrich)
- Urea (Bio-Rad)
- β-Mercaptoethanol (Merck Millipore)

3.2. Instruments / Consumables

- 1.5-ml microtubes (If the objective is to purify protamines for mass spectrometry analyses, it is highly recommended to use low-binding tubes, otherwise it is not required).
- 15-ml tubes
- Calibrated imaging densitometer
- Electrophoretic power supply
- Freezer set at -20ºC
- Fume hood
- Heater set at 37ºC
- Pipettes and tips (10 µl, 20 µl, 200 µl and 1000 µl)
- Platform shaker
- Refrigerated Centrifuge
- Speed-Vacuum
- Vertical mini gel electrophoresis system
- Vortex shaker

4. PROCEDURE

The protocol described in this review is based on the reports from de Yebra and Oliva (1993), de Yebra *et al.* (1993) and Mengual *et al.* (2003) [30,31,34].

4.1. Protamine extraction ● Timing 2h-3h

All solutions should be prepared fresh using ultrapure water (milliQ), and all procedures need to be carried out in ice, unless otherwise noted. ■ **See NOTE 1.**

1. Centrifuge a total amount of $14x10^6$ spermatozoa from each sample to be analyzed at 8940 g for 5 min at 4 ºC. ■ **See NOTE 2.**
2. Wash the pelleted spermatozoa adding 300 µl of either culture medium (*e.g.,* HAM'S F10) or buffer (*e.g.* Pure Sperm® Buffer, or phosphate buffered saline (PBS)). Centrifuge at 8940 g for 5 min at 4 ºC. This step may be repeated in order to ensure the correct elimination of seminal plasma leftovers; however, it is usually not necessary.
3. Suspend the pellet with 200 µl of a solution containing 0.5% Triton X-100, 20 mM (pH8) Trizma® base HCl, and 2 mM $MgCl_2$. Centrifuge at 8940 g for 5 min at 4 ºC and remove the supernatant with a pipette. ■ **See NOTE 3.**
4. Suspend the pellet with 200 µl of 1 mM PMSF. Centrifuge at 8940 g for 5 min at 4 ºC. ■ **See NOTE 4.**

   ► See PROTOCOL MODIFICATION 1
5. Suspend the pellet with 1 volume of a solution composed by 20 mM EDTA, 1 mM PMSF, 100 mM (pH8) Trizma® base HCl. Mix it thoroughly with the pipette. ▼**CRITICAL STEP** Check pH. If pH is not 8, centrifuge at 8940g for 5 min at 4 ºC, and repeat this step as many times as necessary to reach pH 8.
6. Add 1 volume of 575 mM DTT in 6 M guanidinium chloride. ▼**CRITICAL STEP** Mix it thoroughly shaking by hand. It is very important to mix thoroughly or otherwise the chromatin lysis may be incomplete. Do not use pipette since the chromatin denaturation causes the presence of a highly viscous mucus-like gel, which sticks to the pipette tip with potential loss of material. ■ **See NOTE 5.**

   ► See PROTOCOL MODIFICATIONS 2, 3
7. Stop the reaction and precipitate chromatin by adding 5 volumes of cold absolute ethanol. Mix thoroughly by inversion and vortex. It is possible to notice the appearance of a white thread. Incubate at least 10 min at –20 ºC. ■ **See NOTE 6.**

   ▲ PAUSE POINT (If necessary, the sample can be stored at -20ºC for several weeks)
8. Centrifuge at 12880 g for 15 min at 4 ºC and decant to remove the supernatant. Use absorbent paper to eliminate supernatant leftovers remaining in the tube edge.

9. Add 500 µl of 0.5 M HCl. **! CAUTION** Do not use the pipette to mix. Instead mix using vortex and incubate during 5 min at 37 ºC. Subsequently, mix using vortex and incubate once more at 37 ºC for 2 min. Centrifuge at 17530 g for 10 min at 4 ºC. ■ **See NOTE 7.**

   ► **See PROTOCOL MODIFICATION 4**

10. Transfer the supernatant to a pre-cold tube containing enough volume of 100% TCA to obtain a final concentration of 20% TCA and mix thoroughly. ■ **See NOTE 8.**
11. Incubate 10 min at 4 ºC, shaking the tubes every few minutes.
12. Centrifuge at 17530 g for 10 min at 4 ºC. Immediately, remove the supernatant by decantation and use absorbent paper to remove supernatant leftovers in the tube edge. The pellet may not be visible at this step, since the precipitated proteins are bound all along the tube wall.
13. **! CAUTION** At room temperature under the fumehood, wash the proteins bound to the tube wall by adding 500 µl of 1% β-mercaptoethanol diluted in acetone. Centrifuge at 17530 g for 5 min at 4 ºC and discard the supernatant. Perform this step twice. ■ **See NOTE 9.**
14. Dry the proteins in a speed-vacuum for 5-10 minutes with the tube lids opened. Alternatively, allow the proteins to dry overnight under a fume hood at room temperature. Suspend the dried protamines in 20 µl of sample buffer containing 5.5 M urea, 20% β -mercaptoethanol and 5% acetic acid.

    ▲ PAUSE POINT (If necessary, the sample can be stored at -20ºC for several months)

4.2. Acid-urea Polyacrylamide Gel Electrophoresis ● TIMING 6h-7h

All procedures need to be carried out at room temperature, unless otherwise noted.

15. For 2 small polyacrylamide gels (6 cm length and 0.75 mm thickness), prepare 15 ml of a solution composed by 2.5 M urea, 0.9 M acetic acid, and 15% acrylamide/0.1% N,N′-Methylene bis-acrylamide. Degas the solution for 30-40 min under vacuum and subsequently add 80 µl of TEMED and 800 µl of 10% APS. ■ **See NOTE 10.**
16. Add the solution, avoiding bubbles, between the two glasses (previously washed with 70% ethanol and dried) and insert the 15-wells comb.
17. Allow the solution to gelify during 1 h, at room temperature.
18. Remove the comb and wash the wells with distilled H_2O.

4.2.1. Pre-Electrophoresis

19. Place the acid-urea gels in the electrophoresis chamber and fill it with 0.9 N acetic acid buffer. **! CAUTION** Connect the electrodes with the correct polarity at the electrophoresis chamber but with reversed polarity to the power supply. Make sure that the negative electrode is placed at the bottom of the gel. ■ **See NOTE 11.**
20. Run the gel at 150 V during 1h 30 min and monitoring periodically the amperage until it remains constant. ■ **See NOTE 12.**

4.2.2. Electrophoresis

21. Remove the 0.9 N acetic acid buffer and fill the chamber with fresh 0.9 N acetic acid buffer.
22. Wash the wells of the gels with 0.9 N acetic acid buffer using a syringe.
23. Load 2-2.5 µl per well of each sample (*i.e*. protamines in sample buffer, see step 14).
24. Add 3-5 µl of sample buffer with 100 µg/µl methyl green in the first and the last wells to check the electrophoresis progression.
25. Run the gel at 150 V during approximately 55 min. Check that the methyl green mark has progressed three quarters of the gel**!. CAUTION** Connect the electrodes with the correct polarity at the electrophoresis chamber but with reversed polarity at the power supply. Positively charged protamines will migrate to the negatively charge electrode at the bottom of the gel.

4.2.3. EZBlue™ Gel Staining

26. Incubate the gel for 15 min with a solution containing 50% methanol and 10% acetic acid with gentle shaking, in order to fix the proteins in the gel.
27. Incubate the gel for 15 min with milliQ H_2O with gentle shaking, in order to rehydrate the gel.
28. Incubate the gel for at least 1 h with *EZBlue*™, avoiding light exposure, with gentle shaking.
29. Wash the gel with milliQ H_2O for 5 min with gentle shaking. Repeat this step thrice, in order to eliminate staining debris from the gel background.

4.3. Human Protamine P1/P2, Protamine/DNA and Protamine/Sperm Ratios Analyses

For protamine evaluation of the sperm samples, three different ratios can be calculated: P1/P2 ratio, Protamine/DNA ratio, and Protamine/sperm ratio. In any case, the optical densities of the bands corresponding to P1 and P2 family need to be quantified by scanning the gel with an appropriate scanner and using an appropriate software (*e.g*. Quantity One, Bio-Rad). Two main bands are observed in the acid-urea gel: the upper one corresponds to P1, and the lower band to P2 family in human (Figure **2A**). Due to the amino acid composition of P1 combined to its molecular weight, it migrates slower than P2 family.

If protamine quantification is needed, the addition of a protamine standard in each acid-urea gel is required [43, 45] (see protocol modification 1; Figure **2A**). P1/P2 ratios are calculated by either using the protamine optical density measurements or the absolute protamine quantification obtained with the protamine standard. However, the P1/P2 ratio shows two main limitations: (i) it does not provide informa-

tion on whether a specific alteration is due to a deregulation of P1, P2 or both protamines, and (ii) it does not provide information on the overall protamine content in the sperm cell. Alternative ratios have been proposed to solve those issues, such as the Protamine/DNA ratio [45] (see protocol modification 4), and the protamine/sperm ratio [32, 34, 36].

4.4. Protocol Modifications

- **Protocol Modification 1: Histone Removal**

Suspend the pellet obtained after step 4 in 500 µl 0.5 M HCl. Vortex, and incubate for 10 min at 37 ºC. Subsequently, centrifuge at 2000 g for 20 min at 4 ºC. Repeat this step thrice and proceed to step 5. **Note:** The supernatants obtained from the HCl treatments should be stored and used to verify that no protamines have been extracted. Allow histones to precipitate with 20% TCA and visualize by acid-urea gel electrophoresis (Figure **2B**). This protocol modification is recommended for the preparation of a protamine standard necessary for protamine quantification, as described in Mengual *et al.* (2003) [34]. Briefly, the purified protamines after histone removal are quantified by amino acid analysis using an Alpha Plus autoanalyzer (Pharmacia LKB Biotechnology).

- **Protocol Modification 2: Human Protamine Alkylation with Iodoacetamide**

After step 6, add 50 µl of 552 mM iodoacetamide to the mixture, and vortex shortly. Incubate the solution avoiding light exposure at room temperature for 30 min. Proceed to step 7. This protocol modification is recommended to reduce irreversibly the disulfide bonds. However, the iodoacetamide treatment is not usually necessary for human sperm protamine analysis.

- **Protocol Modification 3: Mouse Protamine Alkylation with 4-Vinylpyridine**

After step 6, add 0.8% of 4-vinylpyridine and further incubate the sample for 30 min at 37 ºC, mixing by vortex every 5 min. Proceed to step 7 [81, 82]. This protocol modification is recommended for the extraction of mouse pro-

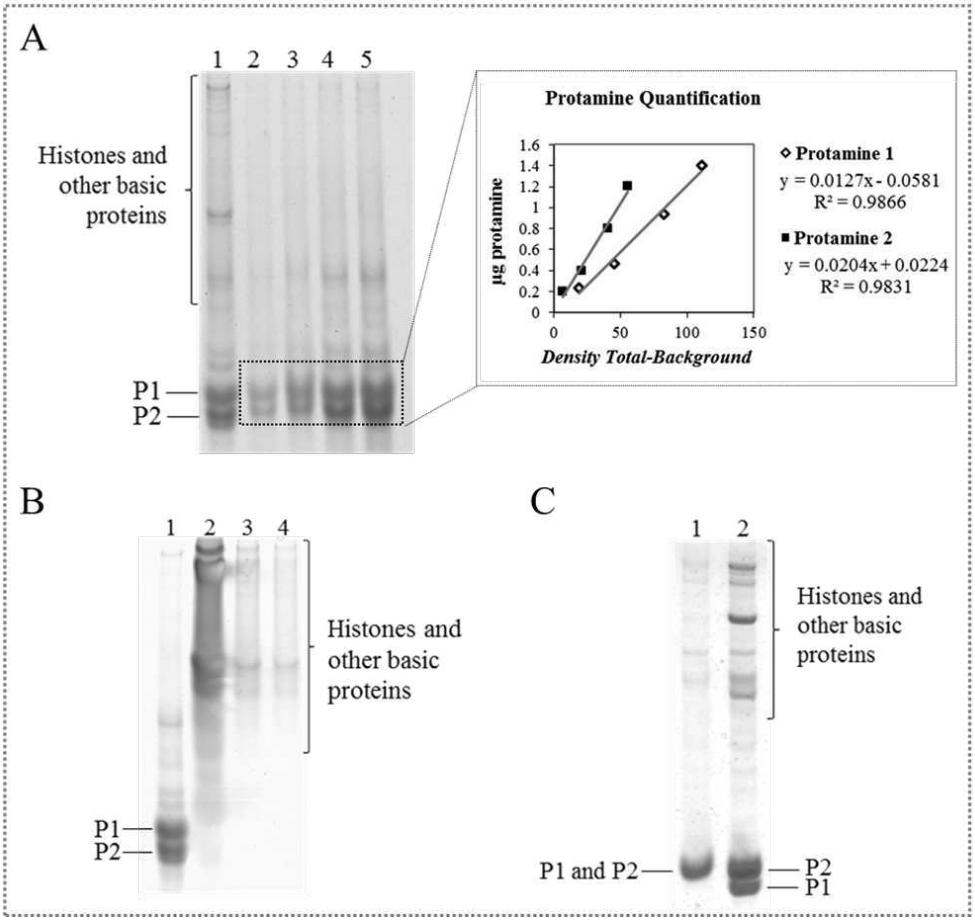


Figure 2. Acid-urea gel electrophoresis of protamines. **A**) Human conventional protamine extraction procedure. Line 1: Protamine 1 (P1) and protamine 2 family (P2) from a fertile male; line 2-5: increasing concentrations of a protamine standard described by Mengual *et al.* (2003). In the right box it is shown the standard curve of P1 and P2 quantification. **B**) Human protamines extraction procedure with histone removal protocol modification. Line 1: P1 and P2 after histone removal; line 2-4: supernatants containing histones. C) Mouse protamines extraction with 4-vinylpyridine protocol modification. Line 1: P1 and P2 without protamine alkylation with 4-vinylpyridine; line 2: P1 and P2 after protamine alkylation with 4-vinylpyridine. Note that the migration of P1 and P2 protamines is reversed when comparing human (A, line 1; B, line 1) with mouse (C, lane 2) [81,82].

tamines. Only one band corresponding to P1 and P2 together is observed in the acid-urea gel electrophoresis, when the regular protamine extraction protocol is followed with mouse samples (Figure **2C**). In order to allow differential electrophoretic migration of mouse P1 and P2, it is required the use of 4-vinylpyridine, which modifies cysteine residues by adding a positive charge (Figure **2C**).

▪ **Protocol Modification 4: DNA Hydrolysis and Quantification**

Use the sediment obtained after step 9 to quantify the amount of DNA of each sample. Briefly, add 500 µl of 0.5 N PCA to the pellet, which contains the DNA. Do not use pipette to mix. Vortex, although the pellet will not be dissolved, and centrifuge at 17530 g for 5 min at 4 °C. Discard the supernatant and add 500 µl of 0.5 N PCA. Submit the pellet to hydrolysis by incubation at 90 °C for 20 min with gentle shaking. After the hydrolysis, place the tubes in ice. Centrifuge at 4420 g for 10 min at 4 °C. Transfer the supernatant to a new tube and determine the absorbance at 260 nm. To quantify the hydrolyzed nucleotides, multiply the absorbance value by the extinction coefficient 38 ng/µl.

This protocol modification is recommended to compare the amounts of P1, P2 and total protamine within samples. After quantification of both protamines and DNA from the same sample, the ratios of P1/DNA, P2/DNA and (P1+P2)/DNA can be calculated.

4.5. Notes

- **NOTE 1:** If the objective is to purify protamines for MS analysis, it is required to use appropriate pure reagents and materials (*e.g.* ultrapure TCA and minimized protein loss tubes).
- **NOTE 2:** The initial amount of sperm may be modified (a range of 5-20 million of sperm per sample is recommended). Nevertheless, and in order to obtain similar efficiencies, the volume of all the solutions necessary for protamine extraction (except for the step 13 of this protocol) need to be adjusted proportionally.
- **NOTE 3:** The application of the detergent Triton X-100 induces sperm membrane permeation, which contributes to the cell lysis in the following step.
- **NOTE 4:** The suspension of the cells in water induces osmotic stress, which causes cell lysis. The PMSF is a protease inhibitor that is incorporated in order to avoid protein degradation. Prepare freshly the PMSF 1 mM solution using a 100 mM PMSF in 2-propanol stock solution (stored in aliquots at -20ºC; before use, aliquots should be warmed to room temperature and verify that no PMSF crystals remain undissolved).
- **NOTE 5:** The application of EDTA, which is a chelating agent with the ability to capture divalent cations; DTT, which is a reducing agent with the ability to rupture disulfide bonds; and guanidine chloride, which is a chaotropic agent with the ability of denaturize proteins, induces DNA denaturation. These steps are necessary for the separation of the sperm chromatin components from other cell fractions that are carried out in the following steps.
- **NOTE 6:** The ethanol treatment allows chromatin precipitation. For a correct precipitation, the samples should be incubated at -20°C for at least 10 min. However, there is no maximum time for this incubation, and the protocol can be paused at this point, leaving the sample at -20°C.
- **NOTE 7:** Basic sperm nuclear proteins are extracted and dissolved in 0.5 M HCl, which allows the separation of the proteins from the DNA. During this step, check that the chromatin and DNA pellet is always in contact with 0.5 M HCl, since it will not be dissolved and it can bind to the tube cap during vortexing.
- **NOTE 8:** The treatment with 20% TCA allows protein precipitation. It is very important to mix thoroughly since 100% TCA stock solution is very dense and if it is not mixed thoroughly it will tend to sink to the bottom of the tube and will remain as a separate stratum.
- **NOTE 9:** If the objective is to purify protamines for intact protein MS analyses, it is highly recommended to use acetone that is suitable for HPLC analysis in this step. Be aware that the use of acetone could lead to the formation of an artifact on the MS spectra. However, TCA precipitation in conjunction with acetone washes would minimize this risk [83]. For the quantification of protamines and protamine PTMs by mass spectrometry, the reader is referred to specific publications [49, 50]
- **NOTE 10:** The 15% acrylamide/0.1% N,N′-Methylene bis-acrylamide solution is prepared by using a solution containing 30% (w/v) acrylamide and 0.2% (w/v) N,N′-Methylene bis-acrylamide, which should be previously filtrated, degassed for 3 hours and stored at 4 °C. Take special precaution, protect yourself with gloves and mask while working with acrylamide, as it is toxic.
- **NOTE 11:** The pre-electrophoresis is required to ensure the elimination of any ions that could disturb the proper development of the electrophoresis. In the subsequent eletrophoresis the electrodes must be connected to the power supply always with negative electrode at the bottom of the gel, in order to allow the top-down migration of the positive proteins.
- **NOTE 12:** During the pre-electrophoresis, the amperage first increases and then starts decreasing until it remains constant (approximately 18 mA for 2 gels). At this point, the correct ions elimination is ensured.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

All authors contributed to draft the manuscript, critically reviewed and approved the final version of the manuscript.

This work was supported by grants to RO from the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competividad; fondos FEDER 'una manera de hacer Europa' PI13/00699, PI16/00346), from Fundación Salud 2000 (SERONO 13-015), and from EUGIN-UB (EU-REP 2014). M.J is granted by Government of Catalonia (Generalitat de Catalunya, pla estratègic de recerca i innovació en salut, PERIS 2016-2020). J.C and A.I are granted by Spanish Ministry of Economy and Competitiveness (Acción estratégica en Salud) CD17/00109 and FI17/00224, respectively. F.B is granted by Spanish Ministry of Education, Culture and Sports (Ministerio de Educación, Cultura y Deporte para la Formación de Profesorado Universitario, FPU15/02306).

REFERENCES

[1] Oliva, R.; Dixon, G.H. Vertebrate protamine genes and the histone-to-protamine replacement reaction. *Prog. Nucleic Acid Res. Mol. Biol.*, **1991**, *40*(C), 25-94.

[2] Balhorn, R.; Corzett, M.; Mazrimas, J.A. Formation of intraprotamine disulfides *in vitro*. *Arch. Biochem. Biophys.*, **1992**, *296*(2), 384-393.

[3] Björndahl, L.; Kvist, U. Human sperm chromatin stabilization: A proposed model including zinc bridges. *Mol. Hum. Reprod.*, **2009**, *16*(1), 23-29.

[4] Kimmins, S.; Sassone-Corsi, P. Chromatin remodelling and epigenetic features of germ cells. *Nature*, **2005**, *434*(7033), 583-589.

[5] Oliva, R. Protamines and male infertility. *Hum. Reprod. Update*, **2006**, *12*(4), 417-435.

[6] Balhorn, R. The protamine family of sperm nuclear proteins. *Genome Biol.*, **2007**, *8*(9), 227.

[7] Rathke, C.; Baarends, W.M.; Awe, S.; Renkawitz-Pohl, R. Chromatin dynamics during spermiogenesis. *Biochim. Biophys. Acta - Gene Regul. Mech.*, **2014**, *1839*(3), 155-168.

[8] Bao, J.; Bedford, M.T. Epigenetic regulation of the histone-to-protamine transition during spermiogenesis. *Reproduction*, **2016**, *151*(5), R55-R70.

[9] Oliva, R.; Castillo, J. Proteomics and the genetics of sperm chromatin condensation. *Asian J. Androl.*, **2011**, *13*(1), 24-30.

[10] Barral, S.; Morozumi, Y.; Tanaka, H.; Montellier, E.; Govin, J.; de Dieuleveult, M.; Charbonnier, G.; Couté, Y.; Puthier, D.; Buchou, T.; Boussouar, F.; Urahama, T.; Fenaille, F.; Curtet, S.; Héry, P.; Fernandez-Nunez, N.; Shiota, H.; Gérard, M.; Rousseaux, S.; Kurumizaka, H.; Khochbin, S. Histone variant H2A.L.2 guides transition protein-dependent protamine assembly in male germ cells. *Mol. Cell*, **2017**, *66*(1), 89-101.e8.

[11] Oliva, R.; Mezquita, C. Marked differences in the ability of distinct protamines to disassemble nucleosomal core particles *in vitro*. *Biochemistry*, **1986**, *25*(21), 6508-6511.

[12] Oliva, R.; Bazett-Jones, D.; Mezquita, C.; Dixon, G.H. Factors affecting nucleosome disassembly by protamines *in vitro*. Histone hyperacetylation and chromatin structure, time dependence, and the size of the sperm nuclear proteins. *J. Biol. Chem.*, **1987**, *262*(35), 17016-17025.

[13] Neto, F.T.L.; Bach, P.V.; Najari, B.B.; Li, P.S.; Goldstein, M. Spermatogenesis in humans and its affecting factors. *Semin. Cell Dev. Biol.*, **2016**, *59*, 10-26.

[14] Jodar, M.; Oliva, R. Protamine alterations in human spermatozoa. *Adv. Exp. Med. Biol.*, **2014**, *791*, 83-102.

[15] Balhorn, R.; Corzett, M.; Mazrimas, J.; Stanker, L.H.; Wyrobek, A. High-performance liquid chromatographic separation and partial characterization of human protamines 1, 2, and 3. *Biotechnol. Appl. Biochem.*, **1987**, *9*(1), 82-88.

[16] Retief, J.D.; Krajewski, C.; Westerman, M.; Dixon, G.H. The evolution of protamine P1 genes in dasyurid marsupials. *J. Mol. Evol.*, **1995**, *41*(5), 549-555.

[17] Queralt, R.; Adroer, R.; Oliva, R.; Winkfein, R.J.; Retief, J.D.; Dixon, G.H. Evolution of protamine P1 genes in mammals. *J. Mol. Evol.*, **1995**, *40*(6), 601-607.

[18] Rooney, A.P.; Zhang, J. Rapid evolution of a primate sperm protein: relaxation of functional constraint or positive darwinian selection? *Mol. Biol. Evol.*, **1999**, *16*(5), 706-710.

[19] Wyckoff, G.J.; Wang, W; Wu, C.I. Rapid evolution of male reproductive genes in the descent of man. *Nature*, **2000**, *403*(6767), 304-309.

[20] Rooney, A.P.; Zhang, J.; Nei, M. An unusual form of purifying selection in a sperm protein. *Mol. Biol. Evol.*, **2000**, *17*(2), 278-283.

[21] Miescher, F. Das protamin eine neue organische base aus den samenfaden des rheinlachses. *Chem. Berlin*, **1874**, *7*, 376-379.

[22] Dahm, R. Friedrich Miescher and the discovery of DNA. *Dev. Biol.*, **2005**, *278*(2), 274-288.

[23] McKay, D.J.; Renaux, B.S.; Dixon, G.H. Human sperm protamines: Amino-acid sequences of two forms of protamine P2. *Eur. J. Biochem.*, **1986**, *156*(1), 5-8.

[24] McKay, D.J.; Renaux, B.S.; Dixon, G.H. The amino acid sequence of human sperm protamine P1. *Biosci. Rep.*, **1985**, *5*(5), 383-391.

[25] Gusse, M.; Sautiere, P.; Belaiche, D.; Martinage, A.; Roux, C.; Dadoune, J.P.; Chevaillier, P. Purification and characterization of nuclear basic proteins of human sperm. *Biochim. Biophys. Acta*, **1986**, *884*(1), 124-134.

[26] Bellve, A.R.; McKay, D.J.; Renaux, B.S.; Dixon, G.H. Purification and characterization of mouse protamines P1 and P2. Amino acid sequence of P2. *Biochemistry*, **1988**, *27*(8), 2890-2897.

[27] Chauvière, M.; Martinage, A.; Debarle, M.; Sautière, P.; Chevaillier, P. Molecular characterization of six intermediate proteins in the processing of mouse protamine P2 precursor. *Eur. J. Biochem.*, **1992**, *204*(2), 759-765.

[28] Pirhonen, A.; Linnala-Kankkunen, A.; Mäenpää, P.H. Identification of phosphoseryl residues in protamines from mature mammalian spermatozoa. *Biol. Reprod.*, **1994**, *50*(5), 981-986.

[29] Chirat, F.; Arkhis, A.; Martinage, A.; Jaquinod, M.; Chevaillier, P.; Sautière, P. Phosphorylation of human sperm protamines HP1 and HP2: Identification of phosphorylation sites. *Biochim. Biophys. Acta*, **1993**, *1203*(1), 109-114.

[30] de Yebra, L.; Oliva, R. Rapid analysis of mammalian sperm nuclear proteins. *Anal. Biochem.*, **1993**, *209*(1), 201-203.

[31] de Yebra, L.; Ballescà, J.L.; Vanrell, J.A.; Bassas, L.; Oliva, R. Complete selective absence of protamine P2 in humans. *J. Biol. Chem.*, **1993**, *268*(14), 10553-10557.

[32] de Yebra, L.; Ballescà, J.L.; Vanrell, J.A.; Corzett, M.; Balhorn, R.; Oliva, R. Detection of P2 precursors in the sperm cells of infertile patients who have reduced protamine P2 Levels. *Fertil. Steril.*, **1998**, *69*(4), 755-759.

[33] Carrell, D.T.; Liu, L. Altered protamine 2 expression is uncommon in donors of known fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. *J. Androl.*, **2001**, *22*(4), 604-610.

[34] Mengual, L.; Ballescà, J.L.; Ascaso, C.; Oliva, R. Marked differences in protamine content and P1 / P2 ratios. *J. Androl.*, **2003**, *24*(3), 438-447.

[35] Nasr-Esfahani, M.H.; Salehi, M.; Razavi, S.; Mardani, M.; Bahramian, H.; Steger, K.; Oreizi, F. Effect of protamine-2 deficiency on ICSI outcome. *Reprod. Biomed. Online*, **2004**, *9*(6), 652-658.

[36] Aoki, V.W.; Liu, L.; Carrell, D.T. Identification and evaluation of a novel sperm protamine abnormality in a population of infertile males. *Hum. Reprod.*, **2005**, *20*(5), 1298-1306.

[37] Aoki, V.W.; Moskovtsev, S.I.; Willis, J.; Liu, L.; Mullen, J.; Brendan M.; Carrell, D.T. DNA integrity is compromised in protamine-deficient human sperm. *J. Androl.*, **2005**, *26*(6), 741-748.

[38] Torregrosa, N.; Domínguez-Fandos, D.; Camejo, M.I.; Shirley, C.R.; Meistrich, M.L.; Ballescà, J.L.; Oliva, R. Protamine 2 precursors, protamine 1/protamine 2 ratio, DNA integrity and other sperm parameters in infertile patients. *Hum. Reprod.*, **2006**, *21*(8), 2084-2089.

[39] Aoki, V.W.; Liu, L.; Jones, K.P.; Hatasaka, H.H.; Gibson, M.; Peterson, C.M.; Carrell, D.T. Sperm protamine 1/protamine 2 ratios are related to in vitro fertilization pregnancy rates and predictive of fertilization ability. *Fertil. Steril.*, **2006**, *86*(5), 1408-1415.

[40] Zini, A.; Gabriel, M.S.; Zhang, X. The histone to protamine ratio in human spermatozoa: comparative study of whole and processed semen. *Fertil. Steril.*, **2007**, *87*(1), 217-219.

[41] Hammadeh, M.E.; Hamad, M.F.; Montenarh, M.; Fischer-Hammadeh, C. Protamine contents and P1/P2 ratio in human spermatozoa from smokers and non-smokers. *Hum. Reprod.*, **2010**, *25*(11), 2708-2720.

[42] de Mateo, S.; Gázquez, C.; Guimerà, M.; Balasch, J.; Meistrich, M.L.; Ballescà, J.L.; Oliva, R. Protamine 2 precursors (pre-P2), protamine 1 to protamine 2 ratio (P1/P2), and assisted reproduction outcome. *Fertil. Steril.*, **2009**, *91*(3), 715-722.

[43] Nanassy, L.; Liu, L.; Griffin, J.; Carrell, D.T. The clinical utility of the protamine 1/protamine 2 ratio in sperm. *Protein Pept. Lett.*, **2011**, *18*(8), 772-777.

[44] García-Peiró, A.; Martínez-Heredia, J.; Oliver-Bonet, M.; Abad, C.; Amengual, M.J.; Navarro, J.; Jones, C.; Coward, K.; Gosálvez, J.; Benet, J. Protamine 1 to protamine 2 ratio correlates with dynamic aspects of DNA fragmentation in human sperm. *Fertil. Steril.*, **2011**, *95*(1), 105-109.

[45] Castillo, J.; Simon, L.; de Mateo, S.; Lewis, S.; Oliva, R. Protamine/DNA ratios and DNA damage in native and density gradient centrifuged sperm from infertile patients. *J. Androl.*, **2011**, *32*(3), 324-332.

[46] Simon, L.; Castillo, J.; Oliva, R.; Lewis, S.E.M. Relationships between human sperm protamines, DNA damage and assisted reproduction outcomes. *Reprod. Biomed. Online*, **2011**, *23*(6), 724-734.

[47] Ni, K.; Spiess, A.N.; Schuppe, H.C.; Steger, K. The impact of sperm protamine deficiency and sperm DNA damage on human male fertility: a systematic review and meta-analysis. *Andrology*, **2016**, *4*(5), 789-799.

[48] Zhang, X.; San Gabriel, M.; Zini, A. Sperm nuclear histone to protamine ratio in fertile and infertile men: Evidence of heterogeneous subpopulations of spermatozoa in the ejaculate. *J. Androl.*, **2006**, *27*(3), 414-420.

[49] Brunner, A.M.; Nanni, P.; Mansuy, I.M. Epigenetic marking of sperm by post-translational modification of histones and protamines. *Epigenetics chromatin*, **2014**, *7*(1), 2.

[50] Castillo, J.; Estanyol, J.M.; Ballescà, J.L.; Oliva, R. Human sperm chromatin epigenetic potential: Genomics, proteomics, and male infertility. *Asian J. Androl.*, **2015**, *17*(4), 601-609.

[51] Corzett, M.; Mazrimas, J.; Balhorn, R. Protamine 1: Protamine 2 stoichiometry in the sperm of eutherian mammals. *Mol. Reprod. Dev.*, **2002**, *61*(4), 519-527.

[52] Hammoud, S.S.; Nix, D.A.; Zhang, H.; Purwar, J.; Carrell, D.T.; Cairns, B.R. Distinctive chromatin in human sperm packages genes for embryo development. *Nature*, **2009**, *460*(7254), 473-478.

[53] Castillo, J.; Amaral, A.; Azpiazu, R.; Vavouri, T.; Estanyol, J.M.; Ballescà, J.L.; Oliva, R. Genomic and proteomic dissection and characterization of the human sperm chromatin. *Mol. Hum. Reprod.*, **2014**, *20*(11), 1041-1053.

[54] Black, J.A.; Dixon, G.H. Evolution of protamine: A further example of partial gene duplication. *Nature*, **1967**, *216*(5111), 152-154.

[55] Bloch, D.P. A catalog of sperm histones. *Genetics*, **1969**, *61*(1), Suppl:93-111.

[56] Subirana, J.A.; Cozcolluela, C.; Palau, J.; Unzeta, M. Protamines and other basic proteins from spermatozoa of molluscs. *Biochim. Biophys. Acta*, **1973**, *317*(2), 364-379.

[57] Kasinsky, H.E.; Eirín-López, J.M.; Ausió, J. Protamines: Structural complexity, evolution and chromatin patterning. *Protein Pept. Lett.*, **2011**, *18*(8), 755-771.

[58] Eirín-López, J.M.; Ausió, J. Origin and evolution of chromosomal sperm proteins. *BioEssays*, **2009**, *31*(10), 1062-1070.

[59] Saperas, N.; Ausió, J. Sperm nuclear basic proteins of tunicates and the origin of protamines. *Biol. Bull.*, **2013**, *224*(3), 127-136.

[60] Luke, L.; Tourmente, M.; Roldan, E.R.S. Sexual selection of protamine 1 in mammals. *Mol. Biol. Evol.*, **2016**, *33*(1), 174-184.

[61] Balhorn, R.; Reed, S.; Tanphaichitr, N. Aberrant protamine 1/protamine 2 ratios in sperm of infertile human males. *Experientia*, **1988**, *44*(1), 52-55.

[62] Lescoat, D.; Colleu, D.; Boujard, D.; Le Lannou, D. Electrophoretic characteristics of nuclear proteins from human spermatozoa. *Arch. Androl.*, **1988**, *20*(1), 35-40.

[63] Bach, O.; Glander, H.J.; Scholz, G.; Schwarz, J. Electrophoretic patterns of spermatozoal nucleoproteins (NP) in fertile men and infertility patients and comparison with NP of somatic cells. *Andrologia*, **1990**, *22*(3), 217-224.

[64] Zatecka, E.; Castillo, J.; Elzeinova, F.; Kubatova, A.; Ded, L.; Peknicova, J.; Oliva, R. The effect of tetrabromobisphenol a on protamine content and DNA integrity in mouse spermatozoa. *Andrology*, **2014**, *2*(6), 910-917.

[65] Scheicher, B.; Lorenzer, C.; Gegenbauer, K.; Partlic, J.; Andreae, F.; Kirsch, A.H.; Rosenkranz, A.R.; Werzer, O.; Zimmer, A. Manufacturing of a secretoneurin drug delivery system with self-assembled protamine nanoparticles by titration. *PLoS One*, **2016**, *11*(11), e0164149.

[66] Tusup, M.; Pascolo, S. Generation of immunostimulating 130 nm protamine-RNA nanoparticles. *Methods Mol. Biol.*, **2017**, *1499*, 155-163.

[67] Lochmann, D.; Stadlhofer, S.; Weyermann, J.; Zimmer, A. New protamine quantification method in microtiter plates using O-phthaldialdehyde/N-acetyl-L-cysteine reagent. *Int. J. Pharm.*, **2004**, *283*(1-2), 11-17.

[68] Stedman, E. Cell specificity of histones. *Nature*, **1950**, *166*(4227), 780-781.

[69] Allfrey, V.G.; Littau, V.C.; Mirsky, A.E. Methods for the purification of thymus nuclei and their application to studies of nuclear protein synthesis. *J. Cell Biol.*, **1964**, *21*, 213-231.

[70] Guillemette, B.; Hammond-Martel, I.; Wurtele, H.; Verreault, A. Production and purification of antibodies against histone modifications. *Methods Mol. Biol.*, **2017**, *1528*, 149-164.

[71] de Mateo, S.; Estanyol, J.M.; Oliva, R. Methods for the analysis of the sperm proteome. *Methods Mol. Biol.*, **2013**, *927*, 411-422.

[72] Liu, L.; Aston, K.I.; Carrell, D.T. Protamine extraction and analysis of human sperm protamine 1/protamine 2 ratio using acid gel electrophoresis. *Methods Mol. Biol.*, **2013**, *927*, 445-450.

[73] Ribas-Maynou, J.; García-Peiró, A.; Martínez-Heredia, J.; Fernández-Encinas, A.; Abad, C.; Amengual, M.J.; Navarro, J.; Benet, J. Nuclear degraded sperm subpopulation is affected by poor chromatin compaction and nuclease activity. *Andrologia*, **2015**, *47*(3), 286-294.

[74] Aoki, V.W.; Emery, B.R.; Liu, L.; Carrell, D.T. Protamine levels vary between individual sperm cells of infertile human males and correlate with viability and DNA integrity. *J. Androl.*, **2006**, *27*(6), 890-898.

[75] Nasr-Esfahani, M.H.; Razavi, S.; Mozdarani, H.; Mardani, M.; Azvagi, H. Relationship between protamine deficiency with fertilization rate and incidence of sperm premature chromosomal condensation post-ICSI. *Andrologia*, **2004**, *36*(3), 95-100.

[76] Khara, K.K. Human protamines and male infertility. *J. Assist. Reprod. Genet.*, **1997**, *14*(5), 282-290.

[77] Simon, L.; Liu, L.; Murphy, K.; Ge, S.; Hotaling, J.; Aston, K.I.; Emery, B.; Carrell, D.T. Comparative analysis of three sperm DNA damage assays and sperm nuclear protein content in couples undergoing assisted reproduction treatment. *Hum. Reprod.*, **2014**, *29*(5), 904-917.

[78] Hamad, M.F.; Shelko, N.; Kartarius, S.; Montenarh, M.; Hammadeh, M.E. Impact of cigarette smoking on histone (H2B) to protamine ratio in human spermatozoa and its relation to sperm parameters. *Andrology*, **2014**, *2*(5), 666-677.

[79] Sakkas, D.; Urner, F.; Bianchi, P.G.; Bizzaro, D.; Wagner, I.; Jaquenoud, N.; Manicardi, G.; Campana, A. Sperm chromatin anomalies can influence decondensation after intracytoplasmic sperm injection. *Hum. Reprod.*, **1996**, *11*(4), 837-843.

[80] Delbes, G.; Hales, B.F.; Robaire, B. Effects of the chemotherapy cocktail used to treat testicular cancer on sperm chromatin integrity. *J. Androl.*, **2007**, *28*(2), 241-249.

[81] Ishibashi, T.; Li, A.; Eirin-Lopez, J.M.; Zhao, M.; Missiaen, K.; Abbott, D.W.; Meistrich, M.; Hendzel, M.J.; Ausió, J. H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis; *Nucleic Acids Res.*, **2010**, *38*(6), 1780-1789.

[82] Crankshaw, M.W.; Grant, G.A. Modification of cysteine. *Curr. Protoc. Protein Sci.*, 2001; *15*, 15.1.1-15.1.18.

[83] Güray, M.Z.; Zheng, S.; Doucette, A.A. Mass spectrometry of intact proteins reveals +98 U chemical artifacts following precipitation in acetone. *J. Proteome Res.*, **2017**, *16*(2), 889-897.

3.3. Treball 3

Caracterització de les proteoformes de les protamines de l'espermatozoide humà mitjançant la combinació de les tècniques d'espectrometria de masses *top-down* i *bottom-up*

Journal of Proteome Research

2020 Jan 3;19(1):221-237

PMID: 31703166

Characterization of Human Sperm Protamine Proteoforms through a Combination of Top-down and Bottom-Up Mass Spectrometry Approaches.

Soler-Ventura A*, Gay M*, Jodar M*, Vilanova M, Castillo J, Arauz-Garofalo G, Villarreal L, Ballescà JL, Vilaseca M, Oliva R.

* Aquests autors han contribuït per igual a aquest treball

Caracterització de les proteoformes de les protamines de l'espermatozoide humà mitjançant la combinació de les tècniques d'espectrometria de masses *top-down* i *bottom-up*

Objectiu: Establir la metodologia per identificar la protamina 1 (P1) i els membres de la família de protamina 2 (P2) i les seves modificacions posttraduccionals (PTMs) associades en mostres independents d'individus normozoospèrmics, per tal de determinar el perfil normal de proteoformes de les protamines.

Metodologia: S'han utilitzat dos tècniques complementàries, la *top-down* espectrometria de masses (MS) i la *bottom-up* MS basada en la digestió amb proteïnasa K de les protamines.

Resultats: La tècnica de *top-down* MS ha permès identificar les formes canòniques de les protamines, és a dir, P1, el precursor de P2 (pre-P2) i les formes madures de P2 (HP2, HP3 i HP4). A més, s'han identificat proteoformes truncades de la pre-P2, incloent tres que ja s'havien descrit prèviament i tres de noves fins ara no descrites. També s'ha descrit una proteoforma truncada per P1, no descrita prèviament. L'aproximació de *bottom-up* MS ha permès identificar pèptids de pre-P2 i confirmar la isoforma 2 de la P2 que no estava anotada a UniProtKB i que també l'hem validat a nivell d'RNA. Pel que fa a les PTMs de les protamines, mitjançant la tècnica de *top-down* MS s'ha identificat la P1 intacta, pre-P2 i les proteoformes madures de P2 i els seus patrons de fosforilació. Per la tècnica de *bottom-up* MS s'han identificat PTMs complementàries al *top-down* MS per la pre-P2. Tanmateix, la isoforma 2 de P2, també ha estat identificada monofosforilada. Inesperadament, a través de la tècnica de *top-down* MS s'ha identificat un increment de +61 Da en diferents proteoformes. Aquesta modificació ha estat identificada sola o en combinació a altres PTMs i podria correspondre a l'addició de Zn^{2+}. La determinació dels ràtios P1/P2 i pre-P2/P2 derivats de dades de *top-down* MS ha permès definir que el ràtio P1/P2 es manté estable entre els diferents pacients analitzats, mentre que el ràtio pre-P2/P2 permet estratificar els pacients en dos subgrups.

Conclusions: En aquest estudi hem demostrat que la combinació de les tècniques de *top-down* i *bottom-up* MS és adequada i dona informació complementària amb una baixa variabilitat interindividual entre els pacients analitzats. A més, s'han identificat proteoformes truncades prèviament descrites i d'altres de noves, tan per P1 com per pre-P2, algunes de les quals contenen residus modificats. Curiosament, a més de la fosforilació, ambdues protamines duen una modificació de +61 Da, que podria correspondre al ió Zn^{2+}. A més, l'enfocament *bottom-up* MS ha permès la identificació de forma inequívoca de la isoforma 2 de P2 no anotada prèviament a UniProtKB. L'avaluació dels ràtios P1/P2 i pre-P2/P2 mitjançant dades derivades de MS en diferents fenotips infèrtils permetran identificar alteracions en proteoformes específiques.

Journal of proteome research
Cite This: J. Proteome Res. 2020, 19, 221−237
Article
pubs.acs.org/jpr

Characterization of Human Sperm Protamine Proteoforms through a Combination of Top-Down and Bottom-Up Mass Spectrometry Approaches

Ada Soler-Ventura,[†,⊥] Marina Gay,[‖,⊥] Meritxell Jodar,[†,⊥] Mar Vilanova,[‖] Judit Castillo,[†] Gianluca Arauz-Garofalo,[‖] Laura Villarreal,[‖] Josep Lluís Ballescà,[§] Marta Vilaseca,[‖] and Rafael Oliva*[,†,‡]

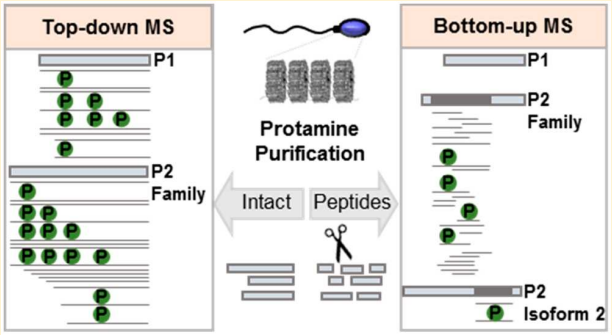
[†]EUGIN-UB Research Excellence Program, Molecular Biology of Reproduction and Development Research Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Fundació Clínic per a la Recerca Biomèdica (FCRB), Faculty of Medicine and Health Sciences, University of Barcelona, 08036 Barcelona, Spain

[‡]EUGIN-UB Research Excellence Program, Biochemistry and Molecular Genetics Service and [§]Clinic Institute of Gynecology, Obstetrics and Neonatology (ICGON), Hospital Clínic, 08036 Barcelona, Spain

[‖]EUGIN-UB Research Excellence Program, Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac, 10, 08028 Barcelona, Spain

Supporting Information

ABSTRACT: Protamine 1 (P1) and protamine 2 (P2) family are extremely basic, sperm-specific proteins, packing 85−95% of the paternal DNA. P1 is synthesized as a mature form, whereas P2 components (HP2, HP3, and HP4) arise from the proteolysis of the precursor (pre-P2). Due to the particular protamine physical−chemical properties, their identification by standardized bottom-up mass spectrometry (MS) strategies is not straightforward. Therefore, the aim of this study was to identify the sperm protamine proteoforms profile, including their post-translational modifications, in normozoospermic individuals using two complementary strategies, a top-down MS approach and a proteinase-K-digestion-based bottom-up MS approach. By top-down MS, described and novel truncated P1 and pre-P2 proteoforms were identified. Intact P1, pre-P2, and P2 mature proteoforms and their phosphorylation pattern were also detected. Additionally, a +61 Da modification in different proteoforms was observed. By the bottom-up MS approach, phosphorylated residues for pre-P2, as well as the new P2 isoform 2, which is not annotated in the UniProtKB database, were revealed. Implementing these strategies in comparative studies of different infertile phenotypes, together with the evaluation of P1/P2 and pre-P2/P2 MS-derived ratios, would permit determining specific alterations in the protamine proteoforms and elucidate the role of phosphorylation/dephosphorylation dynamics in male fertility.



KEYWORDS: *top-down proteomics, bottom-up proteomics, proteoforms, electron transfer dissociation, protamines, sperm, post-translational modifications, phosphorylation, nonannotated proteins, male infertility*

INTRODUCTION

During the final phase of spermatogenesis, called spermiogenesis, human sperm chromatin undergoes a marked remodeling, in which histones are sequentially replaced by specific histone variants, transition proteins, and, finally, protamines.[1] This process results in a unique chromatin structure with 85−95% of sperm DNA packed by protamines and the remaining 5−15% attached to histones.[1−5] Protamines are small and extremely basic sperm-specific proteins rich in positively charged arginine residues (50−70%). This particular amino acid composition promotes the formation of highly condensed toroidal complexes alongside the negatively charged sperm DNA.[1] In addition, protamines harbor cysteine residues able to form disulfide bonds among intra- and interprotamine molecules that stabilize the nucleoprotamine structure.[1,6] It has also been postulated that free thiol groups could bind to metals such as zinc, allowing the formation of salt bridges, which could further have a role in sperm chromatin dynamics.[7] All together this leads to a highly compact chromatin structure that protects sperm DNA from nucleases and contributes in obtaining the required hydrodynamic shape for mature sperm functionality.[1,8−12]

Received: July 24, 2019
Published: November 8, 2019

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

In humans, protamines are encoded by a single copy of protamine 1 (*PRM1*) and protamine 2 (*PRM2*) genes clustered together with the transition nuclear protein 2 (*TNP2*) gene on chromosome 16p13.3.[13,14] Of note, although both protamine genes are present in all mammals, *PRM2* is only translated in some species, including human and mice.[1,15−17] Two types of protamines have been described in human at the protein level, the protamine 1 (P1) and the protamine 2 (P2) family. Interestingly, while P1 is directly synthesized as a mature protein of 51 amino acid residues, P2 mature proteoforms derive from a precursor (pre-P2) of 102 amino acid residues.[18,19] By proteolysis, pre-P2 gives rise to the three P2 mature proteoforms (HP2, HP3, and HP4), comprising residues 46−102 (HP2), 49−102 (HP3), and 45−102 (HP4), differing among them only in the 1−4 amino acid residues from the N-terminus end. HP2 is the most abundant proteoform, followed by HP3 and HP4.[1,16,20−22] Other studies have also identified pre-P2 intermediate processed proteoforms, namely, HPI1, HPI2, HPS1, and HPS2, which could give rise to the mature proteoforms.[23]

The normal ratio between protein amounts of P1 and P2 members was established to be around 1 (0.8−1.2) through acid−urea polyacrylamide gel electrophoresis (PAGE)-based experiments. Alterations in this P1/P2 ratio have been associated with male infertility, sperm DNA damage, and/or impairments in fertilization and embryo development (reviewed elsewhere[24−26]). Nevertheless, Nanassy and colleagues established a normal P1/P2 ratio ranging from 0.54−1.43 in a fertile population.[27] Likewise, contrasting results showing an association between the pre-P2/P2 ratio and an impairment in sperm count, sperm motility, or implantation rate and pregnancy outcome have been reported using PAGE-based methods.[28−31] The presence of different protamine proteoforms, including truncated and/or proteoforms containing different post-translational modifications (PTMs), could explain the variability in these results, since the deregulation of specific proteoforms that could not be detected by standard methods might be the cause of male infertility. Despite the fact that only a small part of sperm DNA remains packed by histones,[1,3,32] sperm histone proteoforms pattern, including testes-specific histone variants and their corresponding PTMs, have been widely studied, and their role as epigenetic marks beyond fertilization has been postulated.[4,33−41] In contrast, the knowledge about protamine proteoforms pattern and their associated potential regulatory activities is scarce.

It is known that during spermatogenesis, P1 and P2 mature proteoforms are extensively phosphorylated by SRSF protein kinase 1 (SRPK1) and calcium/calmodulin dependent protein kinase IV (CAMK4), respectively.[8,42−46] Noteworthy, CAMK4 knockout mice are infertile with spermatogenic defects due to the specific loss of P2 and the long retention of TNP2 in spermatids. These results indicate the importance of protamine phosphorylation in the nucleohistone−nucleoprotamine replacement occurring in the late steps of spermatogenesis.[47] However, once P1 and P2 mature proteoforms are tightly bound to the DNA, an extensive dephosphorylation occurs before spermatozoa enter the epididymis.[46,48] Interestingly, mice deficient in Hspa4l, a heat shock protein and chaperone required for the recruitment of the phosphatase Ppp1cc2 to sperm chromatin, are infertile. This might be most probably due to the low activity of this phosphatase resulting in an abnormal increase of P2 phosphorylation in epididymal sperm.[49] This hypothesis is supported by the rescue of the fertility state of this infertile mouse model when the phosphorylatable P2 residue S56 is substituted by the nonphosphorylatable residue A56.[49] Nowadays few studies based on either phosphoserine conversion followed by protein sequencing or by mass spectrometry (MS) have provided evidence of remaining protamine phosphorylated residues in human mature sperm.[50−52] However, their regular pattern or potential epigenetic role have never been addressed.

In order to perform a comprehensive and accurate characterization of protamine proteoforms in human sperm, advanced high-throughput proteomic techniques are required. Due to the particular physical−chemical properties of protamines (extreme basicity, amino acid sequence rich in arginine and cysteine residues, small size, and high-similarity among P2 mature proteoforms[1]), the use of classical standardized bottom-up MS strategies based on trypsin digestion to arginine and lysine residues is not straightforward.[53] To overcome these pitfalls, technical modifications, such as the use of other proteases, should be optimized. Alternatively, a top-down MS strategy, which is based on intact protein analysis, is also an appropriate option to be considered.[54] Moreover, a top-down MS approach offers unique information over the classic bottom-up strategy, as it can directly identify proteoform microheterogeneity and, thus, report on the combinatorial effects of multiple PTMs, sequence variations, or truncations occurring on the same protein molecule.[55] It has been described that a combination of top-down and bottom-up MS strategies is optimal to describe PTM patterns in proteins with similar characteristics to protamines, for example, the histones.[56] Remarkably, recent studies have applied these two MS approaches to assess protamine PTMs in mouse and human sperm.[52,57] However, all these studies were performed in pooled samples and therefore do not provide information about interindividual variability.

The aim of this study was to set up the methodology to identify P1 and P2 family members and their associated PTMs in independent normozoospermic samples to establish the regular protamine proteoforms profile. Two complementary approaches have been used to accomplish our objective: (1) a top-down MS approach with novel proteomic workflows and data analysis pipelines and (2) a bottom-up MS approach based on proteinase K-digestion of protamines. The implementation of these complementary techniques and the establishment of MS-derived P1 and P2 ratios in future comparative studies including different subtypes of infertile patients could resolve the impact of dysregulations in the protamine proteoforms profile and the importance of their balance in male fertility.

EXPERIMENTAL PROCEDURES

Sample Collection

Human semen samples (n = 10) were obtained from patients undergoing routine semen analysis at the Assisted Reproduction Unit from the Clinic Institute of Gynecology, Obstetrics and Neonatology, at the Hospital Clínic of Barcelona, Spain. All patients gave signed informed consent following the Declaration of Helsinki. The ejaculates were collected by masturbation into sterile containers after 3−5 days of sexual abstinence. The evaluation of the seminal parameters was performed using the automatic semen analysis system CASA (Proiser, Paterna, Spain). After seminal analysis, all samples

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

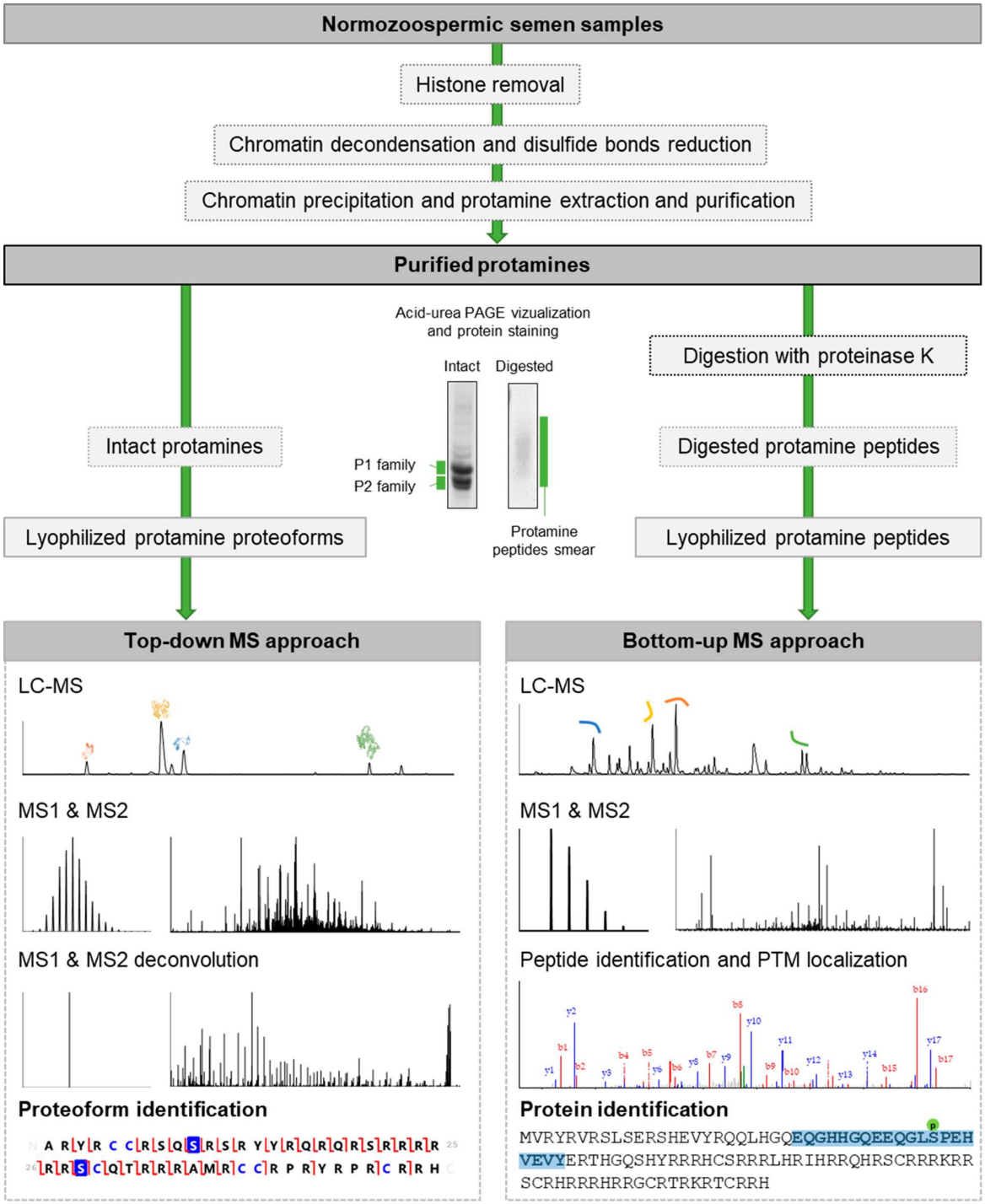


Figure 1. Overall strategy used for the identification of protamines and associated post-translational modifications (PTMs) in human sperm. After removing histones from sperm samples, chromatin was decondensed and disulfide bonds were reduced. Afterward, chromatin was precipitated, and purified protamines were obtained. For the top-down approach, disulfide bonds in intact protamines were rereduced, protected with iodoacetamide, and subjected to liquid chromatography followed by mass spectrometry (LC−MS) after sample clean up. MS1 and MS2 were deconvoluted, leading to the proteoform identification. For the bottom-up approach, purified protamines were digested with proteinase K, and protamine peptides were subjected to LC−MS, which allowed the identification of protamine peptides and their associated PTMs. See Supporting Figure 1 for more details.

were classified as normozoospermic according to World Health Organization guidelines.[58] Sperm samples were purified through 50% Puresperm density gradient separation (NidaCon International AB, Gothenburg, Sweden), following the manufacturer's instructions. Three samples were pooled for the initial bottom-up MS experiments, and one individual sample was employed for bottom-up MS validation. Five samples were used individually for the top-down MS approach

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

and one individual sample was employed for the RNA experiments.

Human Sperm Protamine Purification

Protamine-enriched fraction from spermatozoa was purified as previously described by Soler-Ventura et al. in 2018,[25] except that histones and other basic proteins weakly associated with DNA were removed from the sperm by 0.5 M HCl, vortexing and incubation for 10 min at 37 °C, and centrifugation at 2000*g* for 20 min at 4 °C (three times) prior to suspending the sperm pellet in 0.5% Triton X-100, 20 mM (pH8) Trizma base HCl, and 2 mM $MgCl_2$. Then, samples were processed as described by Soler-Ventura et al.[25] For the top-down MS approach, protamine-enriched fractions equivalent to 15 million spermatozoa were dried at room temperature and kept at −20 °C. Previous to liquid chromatography−mass spectrometry analysis (LC−MS/MS), samples were reconstituted in 50 mM NH_4HCO_3, reduced with 10 mM dithiothreitol (DTT) for 45 min at 56 °C, alkylated for 30 min in the dark with 50 mM iodoacetamide (IAA), and further desalted using PolyLC C_{18} filter tips (PolyLC Inc.). For the bottom-up approach, the protamine-enriched fraction was dried at room temperature and suspended with 20 μL of a solution containing 50 mM Tris-HCl (pH 7.5), 5 mM $CaCl_2$. Subsequently, protamines were digested with proteinase K (Promega, Madison, WI) during 4 h at 37 °C, and the reaction was stopped with 5 mM PMSF. For both the top-down and bottom-up MS approaches, successful protamine purification was monitored by acid−urea PAGE.[25]

Protamine Quantification

Purified extracts of protamines were quantified as described by Soler-Ventura et al.[25] Briefly, purified protamines for each sample were separated by acid−urea PAGE together with increasing amounts of a protamine standard obtained as described elsewhere.[59] Intensities of each band were quantified with Quantity One 1-D analysis software (BioRad, Hercules, CA). The regression curve obtained from increasing amounts of the protamine standard and their band intensities permitted to calculate the approximate quantity of P1 and P2 for each sample, as well as the derived P1/P2 ratio.

Protamine Characterization by High-Throughput Mass Spectrometry-Based Proteomics

The overall procedure used to characterize protamine PTMs is schematized in Figure 1 and the general high-throughput proteomics pipeline is summarized in Supporting Figure 1 of the Supporting Information.

Top-Down Approach. For the top-down MS approach, we employed a total of five biological replicates, and data were integrated. Samples were reconstituted in a 3% acetonitrile (ACN), 1% formic acid (FA) aqueous solution and injected into the LC−MS/MS system.

LC−MS/MS analyses of intact protamines were conducted on two different chromatographic systems [a Dionex Ultimate 3000 (Thermo Scientific, Waltham, MA) or a NanoAcquity Ultra Performance system (Waters Corp., Milford, MA)] coupled to an Orbitrap Fusion Lumos Tribrid Lumos mass spectrometer (Thermo Scientific). Samples were loaded onto C18 trap columns [100 μm × 2 cm Acclaim PepMap100, 5 μm, 100 Å (Thermo Scientific) or 180 μm × 20 cm Symmetry 100 Å, 5 μm (Waters Corp.)] at a flow rate of 15 μL/min. Protein content was eluted and separated using C18 analytical columns [Acclaim PepMap RSLC 75 μm × 50 cm, nanoViper, C18, 2 μm, 100 Å (Thermo Scientific) or BEH300 C18 75 μm × 25 cm, 1.7 μm (Waters Corp.)] using a linear gradient from 3 to 55% solvent B in 60 min at a flow rate of 250 nL/min (solvent A = 0.1% FA in water, solvent B = 0.1% FA in ACN). The Advion TriVersa NanoMate (Advion) was used as the nanospray interface with a spray voltage of 1.60 kV. The mass spectrometer was operated in data-dependent acquisition (DDA) mode. Survey MS scans were acquired on the Orbitrap with the resolution (defined at 200 m/z) set to 120 000. The top N (most intense) ions per scan were fragmented by electron-transfer dissociation (ETD) and detected in the Orbitrap at 120K. The ion count target value was 400 000 for the survey scan and 1 000 000 for the MS/MS scan (ETD). Target ions already selected for MS/MS were dynamically excluded for 30 s. RF lenses were tuned to 30%. The minimal signal required to trigger the MS to MS/MS switch was set to 500 000.

Bottom-Up Approach. For the bottom-up MS approach, an individual sample and the pooled samples (formed by three individual samples) were treated as biological replicates, and the data were integrated. After proteinase K digestion, protamine peptides were diluted in 3% ACN/1% FA and subjected to LC−MS/MS analysis on a Dionex Ultimate 3000 (Thermo Scientific) chromatographic system coupled to an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific) using an Advion nanosource. Sample loading was performed on a 300 μm × 5 mm PepMap100, 5 μm, 100 Å, C18 μ-precolumn (Thermo Scientific) at a flow rate of 15 μL/min and peptides separation on a C18 analytical column [Acclaim PepMap RSLC 75 μm × 50 cm, nanoViper, C18, 2 μm, 100 Å (Thermo Scientific)] at a 250 nL/min flow rate. A linear gradient from 1 to 35% solvent B in 60 min was used (solvent A = 0.1% FA in water, solvent B = 0.1% FA in ACN). Survey MS scans were acquired on the Orbitrap with the resolution (defined at 200 m/z) set to 120 000. The top-speed (most intense) ions per scan were fragmented by high-collision dissociation (HCD) and detected by the Orbitrap with a resolution set to 30 000. The ion count target value was 400 000 and 10 000 for the survey scan and the MS/MS scan, respectively. Target ions already selected for MS/MS were dynamically excluded for 15 s.

Proteomics Data Analysis

The mass spectrometry data have been deposited to the PRIDE Archive (http://www.ebi.ac.uk/pride/archive/) via the PRIDE partner repository with the data set identifier PXD014618.

Top-Down Approach. To perform the MS data analysis for proteoform identification and subsequent PTM location, three different software/search node combinations were employed: BioPharma Finder, ProSight PD, and TopPic. The resulting search outputs were combined in a single harmonized data set and the experimental deconvoluted masses were classified using the density-based spatial clustering of applications with noise (DBSCAN) algorithm.[60] Finally, we assigned to each mass cluster a common proteoform leading to the final proteoform characterization.

BioPharma Finder v 1.0 SP1 (Thermo Fisher Scientific) was used to extract averaged mass spectra from detected chromatographic peaks. Deconvolution was done using the auto Xtract algorithm on resolved m/z charged species.[61,62] The slice window option was set on with a target average spectrum width of 0.1 min. Protamine sequences were introduced to find

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

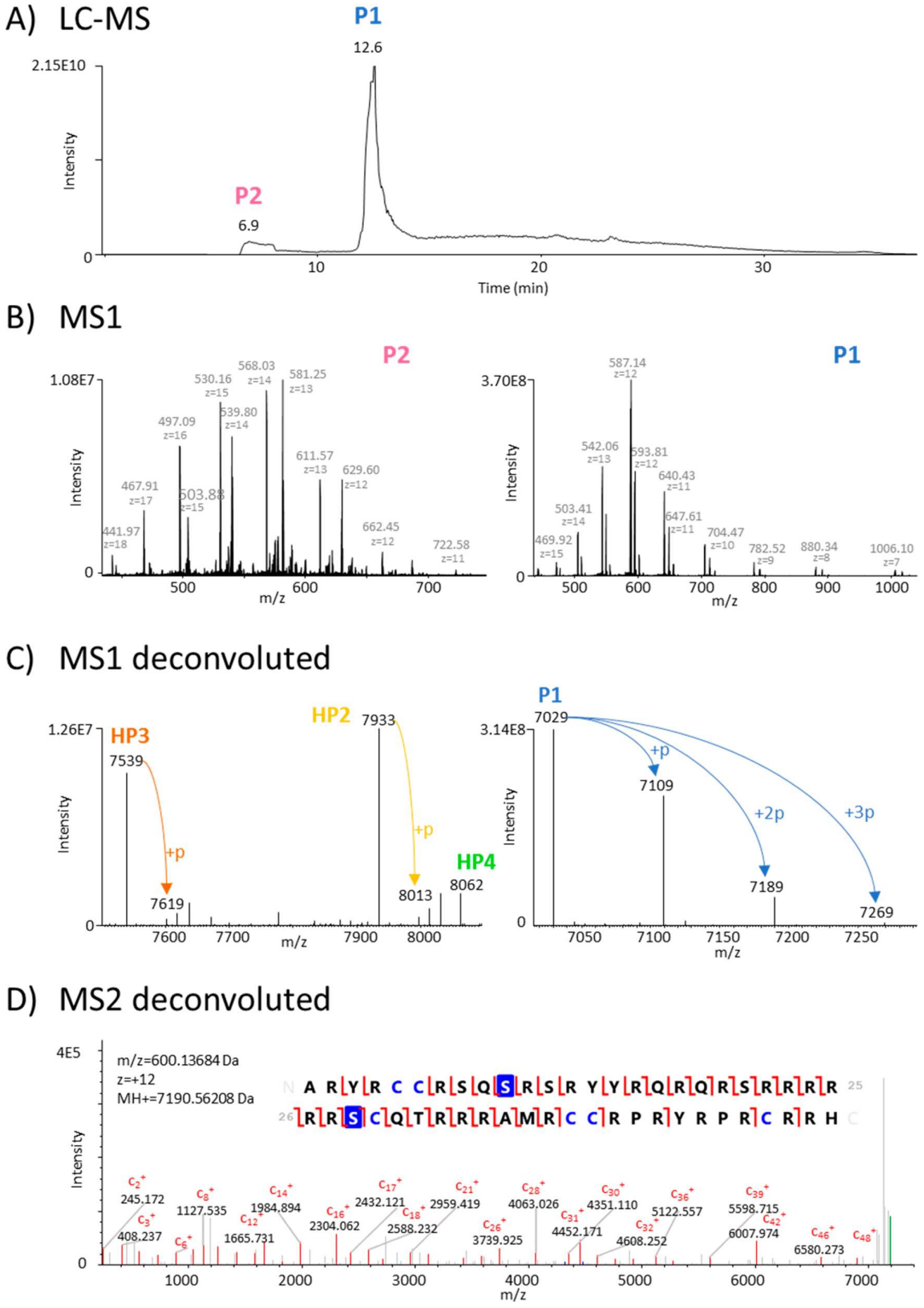


Figure 2. Overall top-down mass spectrometry workflow. (A) Chromatogram resolving protamine 2 (P2) and protamine 1 (P1) at different elution times. (B) Full MS1 spectra of P2 and P1. (C) Deconvolution of P2 with its mature forms unmodified HP3 and with one phosphorylation, unmodified HP2 and with one phosphorylation, and unmodified HP4, and P1 with the phosphorylation pattern. (D) MS2 deconvoluted spectrum for P1 + 2 × ph and its fragment map showing all detected *c* and *z* ions and p-site localizations (phosphorylated amino acids are marked in blue).

a match between the experimental and theoretical masses. Proteome Discoverer v 2.1.0.81 (Thermo Fisher Scientific) with ProSightPD v 1.1 node[63] and TopPic Suite v 1.1.1[64] were used for further protamine identification and PTM assignment. Database searches in ProSight were performed using an xml-annotated database from SwissProt, containing P1 and P2

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

entries. The annotated database was modified to include phosphorylation modification to all STY residues, acetylation in KS, methylation in KR, and dimethylation in KR. A three-tier search (Absolute, Biomarker, and Absolute) was performed with precursor mass tolerances of 2 Da/10 ppm/500 Da, respectively, and fragment mass tolerances of 10 ppm. Searches in TopPic were done using the TopFD tool for top-down spectral deconvolution and a FASTA format database, which contained P1 and P2 sequences. Common modifications were included (oxidations in M, phosphorylation in STY, acetylation in KS, methylation in KR, and dimethylation in KR). Precursor and fragment mass tolerances were set to 15 ppm, and maximum mass shift was set to 500 Da. Proteoform spectrum matches with FDR < 1% were considered for further data integration and clustering analysis.

BioPharma Finder, ProSight, and TopPic search outputs were integrated into a single data set, which unified all main magnitudes provided by each software (retention time, deconvoluted mass, intensity, scan number, proteoform sequence, etc.). Then, the scale of experimental deconvoluted masses was standardized, and the resulting samples were classified through a DBSCAN algorithm with 0.004 as the maximum sample distance and 3 as the sample minimum number (detailed interpretation of the two DBSCAN algorithm parameters is provided in ref 60). This clustering algorithm is particularly convenient for large top-down proteomics data sets with an arbitrary number of proteoforms because is fast and it handles data sets with arbitrary cluster number. After annotating the protein associated with each cluster, the difference between the mass of the corresponding unmodified protein and the average deconvoluted mass of each cluster was computed. The resulting delta mass finally allowed annotating the proteoform associated with each cluster. Additionally, some of the proteoform PTM sites were further validated using ProSight Lite.[65] The average and individual intensities for each proteoform and biological replicate were calculated using data from BioPharma Finder and TopPic outputs, since those correspond to the deconvoluted MS1 spectrum and LC−MS features, respectively, and correlate with the number of proteoforms. In contrast, ProSight data corresponds to MS2 intensities, which are not related to the amount of a specific proteoform. MS $\log_2$(intensities) data of P1, P2, and pre-P2 proteoforms were also used to calculate P1/P2 and pre-P2/P2 ratios for each of the biological replicates.

Bottom-Up MS Approach. Database searches were performed with Proteome Discoverer software v 2.3.0.480 (Thermo Scientific) with Sequest HT search engine and PEAKS Studio software v 8.5 (Bioinformatic Solutions Inc., Waterloo, Canada) using the UniProtKB database [human, release 2018_01 and contaminants]. Searches were run against targeted and decoy databases to determine the false discovery rate (FDR). Search parameters included no enzyme specificity. Oxidation in M, phosphorylation in STY, and acetylation in protein N-terminus were considered as dynamic modifications. Peptide mass tolerance was set to 10 ppm and the MS/MS tolerance was 0.02 Da. Peptides with a FDR < 1% were considered as positive identifications with a high confidence level. Peptides were filtered using a 0.4% FDR, ALC ≥ 50%, and PTM A-score > 20.

Confirmation of P2 Isoform 2 at the RNA Level

Purified spermatozoa from an ejaculate were prepared and stored at −80 °C in RLT buffer (Qiagen, Hilden, Germany) and 1.5% β-mercaptoethanol until further processing. Sperm long-RNA fraction was extracted following the RNeasy Mini Kit protocol (Qiagen) with some modifications previously established by Goodrich et al. and Jodar et al.[66,67] High-quality sperm RNA free of DNA and somatic RNA contamination[66,67] was employed for further analysis (Primers used for quality controls are shown in Supporting Table 1). To validate the presence of P2 isoform 2 at the RNA level, 50 ng of sperm RNA was reverse-transcribed using Maxima RT (Thermo Fisher Scientific) and random primers (Applied Biosystems), followed by real-time PCR using PowerUp SYBR Green PCR (Applied Biosystems) targeting a specific region of the P2 isoform 2 transcript (Supporting Table 1 and Supporting Figure 2). PCR products were visualized in a 3% agarose gel at 90 V. After that, two bands were obtained, one showing the P2 isoform 2 predicted size and another one corresponding to a product shorter in length (Supporting Figure 2). Both bands were excised and eluted using the QIAEX II Gel Extraction Kit (Invitrogen) according to the manufacturer's instructions. The new product was directly analyzed by direct sequencing (Supporting Figure 2), whereas P2 isoform 2 required seminested PCR using PowerUp SYBR Green real-time PCR (Thermo Fisher Scientific) with specific primers (Supporting Table 1) before direct sequencing (Supporting Figure 2).

RESULTS AND DISCUSSION

The LC for the top-down MS strategy was able to separate P2 from P1 at different elution times (Figure 2). Subsequent MS/MS analysis identified 45 different proteoforms, and comparison of the five distinct biological replicates revealed that, in terms of intensity, most abundant proteoforms were detected in all biological replicates. Remarkably, interindividual variability was quite low, since 60% of the proteoforms were identified in all the individuals or in four out of five samples (40% and 20%, respectively), 22.2% in three out of five samples, 11.11% in two out of five samples, and only 6.67% appeared in a unique biological replicate (Table 1 and Supporting Figure 3). These results highlighted that almost 82% of the identified proteoforms was present in more than half of the studied individuals, revealing reproducible high-quality data acquirement. Using this approach, we were able to identify intact unmodified P1; pre-P2 and P2 mature proteoforms; previously described pre-P2 truncated proteoforms and two novel truncated proteoforms for both P1 and pre-P2; and different combinations of PTMs for P1, pre-P2, HP2, and HP3 (Table 1, Supporting Table 2, Figure 3, and Supporting Figures 3−6).

Additionally, with the bottom-up MS workflow, we detected a total of 156 peptide spectrum matches (PSMs) comprising 57 different peptides. Particularly, the detected peptides corresponded to the intact unmodified pre-P2, as well as to the mono- and di- phosphorylated pre-P2 forms. In addition, peptides corresponding to the unmodified P2 isoform 2 and its monophosphorylated proteoform were also found (Table 2 and Figure 4). Approximately 40% of the peptides were identified in both the pooled samples and the individual sample (Table 2). Of note, P1 and P2 mature proteoforms were not detected by the bottom-up MS approach, probably due to either the small size of P1 peptides after proteinase K digestion or the high similarity among the sequences of P2 mature components, respectively.

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

Table 1. Protamine Proteoforms Identified through the Top-Down MS Approach

protein	proteoform[a]	mass (Da) theoretical	mass (Da) experimental	error (Da)	experimental mass shift (Da)	total counts	total average $\log_2$	$\log_2$(intensity) b1	b2	b3	b4	b5
P1	P1	7029.596	7029.9 ± 0.6	−0.283	0.0 ± 0.6	3812	32.34	33.35	32.34	30.89	30.97	31.23
P1	P1 + 61		7090.9 ± 0.7		61.0 ± 0.7	2868	32.00	31.69	30.75	32.90	31.28	32.49
P1	P1 + ph	7109.562	7109.7 ± 0.7	−0.161	79.8 ± 0.7	1343	31.90	33.26	30.70	28.95	30.38	31.04
P1	P1 + 2 × 61		7151.8 ± 1.1		122.0 ± 1.1	1217	30.85	28.99	28.02	32.26	29.97	30.81
P1	P1 + ph + 61		7170.7 ± 0.8		140.9 ± 0.8	1294	31.73	31.30	29.96	32.56	31.70	32.82
P1	P1 + 2 × ph	7189.528	7189.3 ± 1.5	0.278	159.4 ± 1.5	147	29.72	32.66	25.28	24.72	25.44	26.24
P1	P1 + 3 × 61		7212.8 ± 0.9		183.0 ± 0.9	264	29.51	26.79		30.50	28.02	26.93
P1	P1 + ph + 2 × 61		7231.9 ± 0.7		202.0 ± 0.7	951	30.95	29.64	25.52	33.27	30.08	30.55
P1	P1 + 2 × ph + 61		7250.7 ± 0.8		220.8 ± 0.8	138	29.88	28.59	25.25	28.59	32.05	25.47
P1	P1 + 3 × ph	7269.495	7270 ± 4	−0.532	240 ± 4	11	25.48	25.67	25.63	25.57		24.34
P1	P1 + ph +3 × 61		7292.8 ± 0.6		263.0 ± 0.6	124	28.90	26.72		29.75	28.02	28.65
P1	P1 + 2 × ph + 2 × 61		7311.5 ± 1.2		281.6 ± 1.2	111	26.91	25.07	23.15	29.18	27.04	27.95
P1	P1 + 3 × ph + 61		7330.5 ± 0.4		300.6 ± 0.4	19	25.35		23.91	24.35	26.34	26.80
P1	P1 + 2 × ph + 3 × 61		7372.8 ± 0.4		342.9 ± 0.4	7	24.38			24.91	23.41	23.91
P1 [8 − 51]	P1 [8 − 51]	6163.232	6163.257 ± 0.01	−0.026	0.000 ± 0.016	9	27.68	29.39	24.88			
P1 [8 − 51]	P1 [8 − 51] + ph	6243.198	6243.238 ± 0.008	−0.039	79.980 ± 0.008	3	27.16	27.16				
P2	P2	13196.820	13196.8 ± 1.2	0.066	0.0 ± 1.2	178	29.95		31.45	28.05	29.61	
P2	P2 + ph	13276.786	13276.5 ± 1.5	0.273	79.8 ± 1.5	77	27.27		27.76	27.37	26.59	
P2	P2 + 2 × ph	13356.753	13356.7 ± 1.6	0.090	159.9 ± 1.6	77	26.97		25.71	27.29	27.11	
P2	P2 + 3 × ph	13436.719	13436.6 ± 1.3	0.110	239.9 ± 1.3	48	25.56		24.48	26.20	25.78	
P2	P2 + 4 × ph	13516.685	13515.8 ± 1.7	0.905	319.9 ± 1.7	27	25.51		27.06	20.67	25.78	
P2 [8 − 102]	P2 [8 − 102]	12367.316	12367.5 ± 0.9	−0.163	0.0 ± 0.9	7	24.30		25.52		23.09	
HPI2	HPI2	10654.457	10654.5 ± 0.7	−0.009	0.0 ± 0.7	24	26.89		26.70	27.42	26.69	
HPS1	HPS1 − Q	9172.854	9172.8 ± 0.5	0.027	−128.8 ± 0.05	70	27.54		27.13	28.03	27.32	26.85
HPS1	HPS1 − Q + 61		9233.6 ± 0.6		−68.1 ± 0.6	3	25.69			25.69		
HPS1	HPS1 − Q + ph	9252.820	9252.7 ± 1.0	0.142	−48.9 ± 1.0	6	25.64			26.31		24.97
HPS1	HPS1 + pyroglu	9283.886	9284.0 ± 0.9	−0.145	−17.6 ± 0.9	75	27.05		28.28	25.92	27.29	24.04
HPS1	HPS1	9300.912	9301.6 ± 0.9	−0.704	0.0 ± 0.9	360	28.82	29.33	29.32	28.79	28.78	27.42
HPS1	HPS1 + pyroglu + 61		9344.0 ± 1.5		42.4 ± 1.5	8	24.95		*[b]	24.95		
HPS1	HPS1 + 61		9363.5 ± 1.1		61.9 ± 1.1	77	27.53		25.79	28.24	26.26	23.36
HPS1	HPS1 + ph	9380.879	9381.8 ± 0.4	−0.941	80.2 ± 0.4	16	25.82	27.14	27.22	24.12	26.48	23.92
HPS2	HPS2	9002.748	9002.7 ± 0.8	0.022	0.0 ± 0.8	459	29.56	29.90	29.95	29.11	28.70	30.08
HPS2	HPS2 + 61		9064 ± 2		61 ± 2	43	27.45	24.22	27.85	27.63	27.49	*[b]
P2 [38 − 102]	P2 [38 − 102]	8915.716	8916.0 ± 0.5	−0.264	0.0 ± 0.5	4	26.98				26.98	
P2 [39 − 102]	P2 [39 − 102]	8818.663	8818.4 ± 1.2	0.228	0.0 ± 1.2	8	25.51			26.36	23.86	25.47
HP2	HP2	7933.277	7933.2 ± 0.5	0.058	0.0 ± 0.5	1663	31.53	32.03	30.85	31.18	31.91	31.36
HP2	HP2 + 61		7994.3 ± 0.7		61.0 ± 0.7	416	30.85	30.88	27.08	31.33	31.51	31.42
HP2	HP2 + ph	8013.243	8013.3 ± 0.4	−0.030	80.1 ± 0.4	15	24.80	26.35	23.81			23.58
HP2	HP2 + 2 × ph	8093.209	8092.282 ± 0.002	0.927	159.064 ± 0.002	4	25.21	24.48		25.83	25.09	25.46
HP3	HP3	7539.069	7539.1 ± 0.6	0.009	0.0 ± 0.6	929	30.65	30.98	29.74	30.75	30.75	30.65
HP3	HP3 + 61		7600.1 ± 0.5		61.6 ± 0.5	145	29.98	29.37	25.89	30.12	29.78	30.98
HP3	HP3 + ph	7619.035	7619.06 ± 0.02	−0.021	80.00 ± 0.002	3	23.76	24.13	23.38			

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

Table 1. continued

protein	proteoform[a]	mass (Da) theoretical	mass (Da) experimental	error (Da)	experimental mass shift (Da)	total counts	total average $\log_2$	$\log_2$(intensity) b1	b2	b3	b4	b5
HP3	HP3 + 2 × ph	7699.002	7698.074 ± 0.003	0.927	159.014 ± 0.003	3	24.26			24.47	23.87	24.43
HP4	HP4	8062.319	8061 ± 2	0.461	0 ± 2	97	27.02	28.26	26.63	25.87	25.34	25.39
HP4	HP4 + 61		8122.7 ± 1.4		60.9 ± 1.4	5	24.57	24.26		25.04	24.03	24.96

[a]Abbreviations: ph = phosphorylated, pyroglu = pyroglutamic acid modified. [b]Proteoforms identified only with ProSight, which does not provide intensity information.

Novel Protamine Proteoforms

Using the top-down MS approach, the canonical protamines, namely, P1 (P04553), pre-P2 (P04554), and P2 mature proteoforms HP2, HP3, and HP4, were fully identified (Table 1 and Figure 3). Moreover, we identified previously described pre-P2 truncated proteoforms, spanning residues 22−102 (HPI2), 34−102 (HPS1), and 37−102 (HPS2).[68−70] Interestingly, we detected four novel truncated proteoforms, one spanning residues 8−51 of P1 and the others spanning residues 8−102, 38−102, and 39−102 of pre-P2 (Table 1, Figure 3, and Supporting Figure 6). Although it has been proposed that pre-P2 is highly proteolytically processed in a regulated manner,[71,72] the mechanisms and the specific team players implicated in this processing have not yet been described.[23] In accordance with pre-P2 processing, similar post-translational cleavage could proteolyze the truncated form of P1. In both cases, the functions and roles of these truncated forms remain to be clarified.

With the bottom-up MS approach, we identified peptides corresponding to pre-P2 and P2 isoform 2 (P04454-2), which is not annotated and was only validated at the RNA level according to the UniProtKB database (https://www.uniprot.org/).[73] Here, we identified with clear spectra two unique peptides of P2 isoform 2 with 9 and 10 amino acid residues, respectively, after proteinase K digestion (Table 2 and Figure 4). Remarkably, a unique peptide of P2 isoform 2 was found annotated at Nextprot from 2016.[74] Therefore, our findings together with that previous evidence fully confirm the presence of this P2 isoform 2 in human sperm at the protein level.

P2 isoform 2 arises from the splice site variant ENST00000-435245.2 and results in a 140 amino acid length protein differing from pre-P2 by residue 91.[73,75] The presence of P2 isoform 2 was also validated here at the RNA level in an independent sperm sample. Surprisingly, two different products were amplified using a specific set of primers for P2 isoform 2 (Figure 5). Direct sequencing of each product revealed two splice variants of P2, the already known P2 isoform 2 and a novel splice variant that we named here as P2 isoform 3 (Figure 5 and Supporting Figure 2). However, P2 isoform 3 was not detected at the protein level using either top-down or bottom-up MS approaches, suggesting that either this product is not translated or the employed techniques were unable to identify it at the protein level. Looking in detail at the intron nucleotide sequence of *PRM2*, the splice sites corresponding to P2 isoform 2 and isoform 3 were solely conserved among different higher primate species but not in other mammals, such as *Mus musculus* and *Rattus norvegicus*, suggesting that there is no specific evolutionary conservation of these isoforms in mammals (Supporting Figure 7). Further studies are necessary to elucidate the potential role of these protamine isoforms.

Protamines Contain Multiple Post-Translational Modifications

One of the biggest challenges in top-down proteomics studies is PTM localization. Despite the better MS/MS protein fragmentation efficiency with current instrumentation and the improvements in top-down data analysis software, which includes the identification of unknown PTMs, there is still a lack of reliable PTM localization assignment. ProSight gives a C-score related with proteoform characterization[76] and TopPic provides a PTM localization probability. PTM automatic localization information was employed to get the phosphor-

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

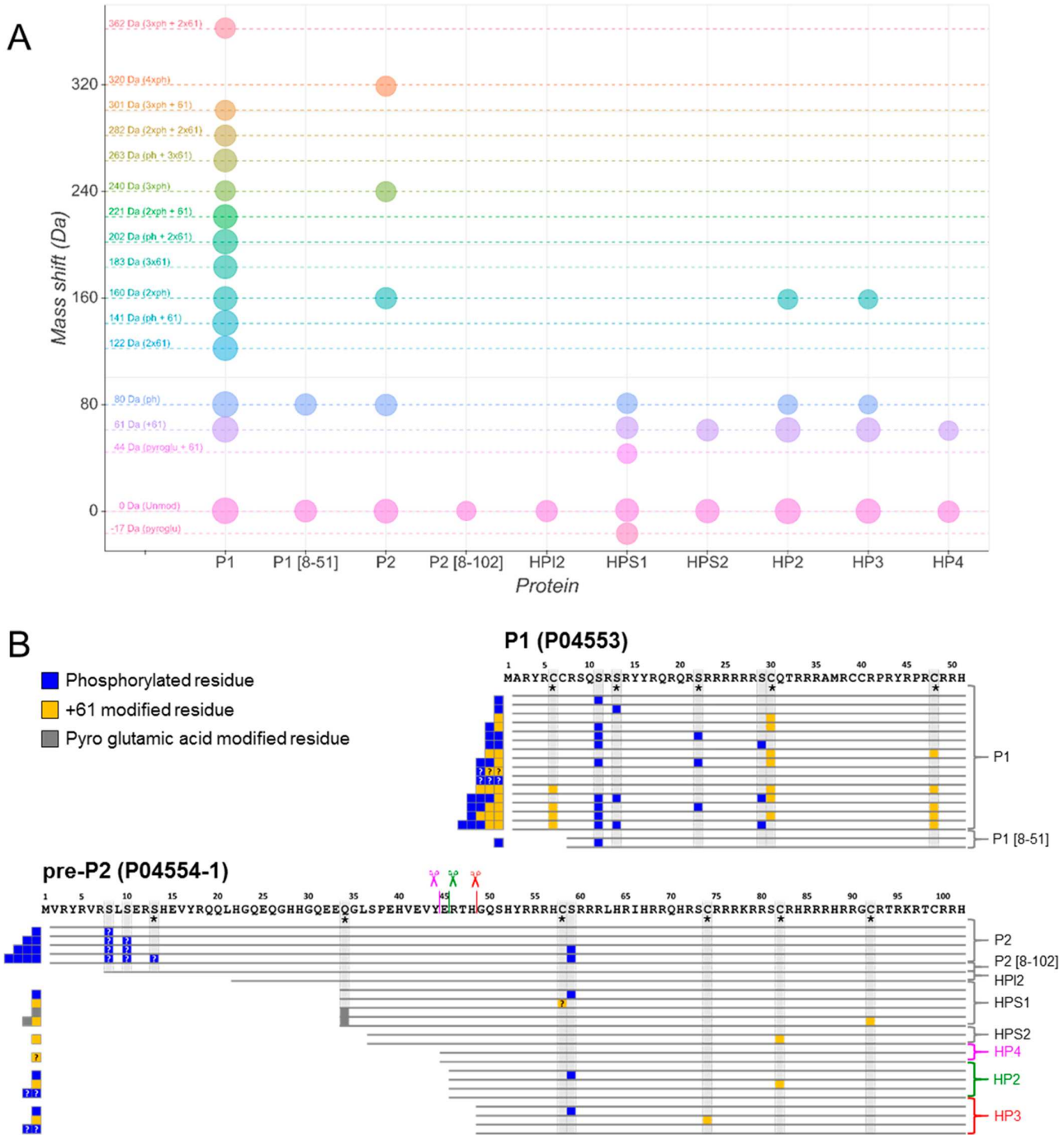


Figure 3. Protamine 1 (P1) and protamine 2 (P2) proteoform profile obtained by our top-down mass spectrometry approach. (A) Scatter plot of the experimental mass shifts [Mass shift (Da)] versus the protamine family proteins (Protein). Thus, each dot represents a distinct protamine proteoform (regardless the PTM localization site), their size being proportional to each proteoform relative intensity on a $\log_2$ scale. The different sets of PTM combinations are highlighted by using a sequential color scale in increasing mass shift order. (B) Schematic representation of P1 and P2 proteoforms identified with their associated PTMs localizations. Phosphorylated, +61 Da, and pyroglutamic acid modified residues are indicated in blue, orange, and gray, respectively. Scissors indicate the proteolytic site for mature forms (HP2 in green, HP3 in red, and HP4 in pink). Question marks at the left indicate those PTMs without fragmentation information available. Question marks within the sequence schemes denote those PTMs with not enough fragment pattern information to accurately localize them. Asterisks mark novel modified residues.

ylation localization in P1 and P2 proteoforms (Supporting Figure 8). Nevertheless, TopPic only provides a PTM localization probability for monomodified proteoforms of known PTMs, and ProSight does not return a site localization probability. In this regard, according to our cluster analysis (Figure 3A), few identified proteoforms were successfully site-localized by ProSight or TopPic. In order to validate and complete the protamine proteoforms map, ProSight Lite[65] was

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

Table 2. Protamine Peptides Identified with the Bottom-Up MS Approach[a]

no. of amino acid residues					PSMs	
initial	final	sequence	modification	amino acid residue modified	pool	rep 1
pre-P2 (P04554)						
8	17	R.SLSERSHEVY.R	phosphorylation	S13	1	
8	17	R.SLSERSHEVY.R			1	
8	19	R.SLSERSHEVYRQ.Q			1	1
8	20	R.SLSERSHEVYRQQ.L			1	1
20	44	Q.QLHGQEQGHHGQEEQGLSPEHVEVY.E				1
21	44	Q.LHGQEQGHHGQEEQGLSPEHVEVY.E			1	2
22	44	L.HGQEQGHHGQEEQGLSPEHVEVY.E			2	2
23	44	H.GQEQGHHGQEEQGLSPEHVEVY.E	phosphorylation	S37	1	
23	44	H.GQEQGHHGQEEQGLSPEHVEVY.E			1	2
25	39	Q.EQGHHGQEEQGLSPE.H				1
25	42	Q.EQGHHGQEEQGLSPEHVE.V			1	
25	44	Q.EQGHHGQEEQGLSPEHVEVY.E	phosphorylation	S37	1	1
25	44	Q.EQGHHGQEEQGLSPEHVEVY.E			4	3
27	39	Q.GHHGQEEQGLSPE.H	phosphorylation	S37		1
27	39	Q.GHHGQEEQGLSPE.H				1
27	42	Q.GHHGQEEQGLSPEHVE.V			2	1
27	43	Q.GHHGQEEQGLSPEHVEV.Y			1	
27	44	Q.GHHGQEEQGLSPEHVEVY.E	phosphorylation	S37	3	1
27	44	Q.GHHGQEEQGLSPEHVEVY.E			4	3
28	44	G.HHGQEEQGLSPEHVEVY.E	phosphorylation	S37	1	
28	44	G.HHGQEEQGLSPEHVEVY.E			1	1
29	42	H.HGQEEQGLSPEHVE.V			1	
29	44	H.HGQEEQGLSPEHVEVY.E	phosphorylation	S37	1	
29	44	H.HGQEEQGLSPEHVEVY.E			2	
30	39	H.GQEEQGLSPE.H			2	1
30	42	H.GQEEQGLSPEHVE.V			1	
30	44	H.GQEEQGLSPEHVEVY.E	phosphorylation	S37	1	
30	44	H.GQEEQGLSPEHVEVY.E			2	1
31	39	G.QEEQGLSPE.H			1	
31	42	G.QEEQGLSPEHVE.V			1	
31	43	G.QEEQGLSPEHVEV.Y			1	
31	44	G.QEEQGLSPEHVEVY.E	phosphorylation	S37	2	1
31	44	G.QEEQGLSPEHVEVY.E			3	1
32	41	Q.EEQGLSPEHV.E			1	
32	42	Q.EEQGLSPEHVE.V	phosphorylation	S37		1
32	42	Q.EEQGLSPEHVE.V			1	
32	43	Q.EEQGLSPEHVEV.Y			1	
32	44	Q.EEQGLSPEHVEVY.E	phosphorylation	S37	3	
32	44	Q.EEQGLSPEHVEVY.E			5	1
33	44	E.EQGLSPEHVEVY.E	phosphorylation	S37	1	
33	44	E.EQGLSPEHVEVY.E			1	1
34	42	E.QGLSPEHVE.V			1	
34	44	E.QGLSPEHVEVY.E			1	
35	43	Q.GLSPEHVEV.Y			1	
35	44	Q.GLSPEHVEVY.E	phosphorylation	S37	3	1
35	44	Q.GLSPEHVEVY.E			8	1
36	44	G.LSPEHVEVY.E			2	1
37	43	L.SPEHVEV.Y				3
37	44	L.SPEHVEVY.E	phosphorylation	S37	1	
37	44	L.SPEHVEVY.E			31	3
37	46	L.SPEHVEVYER.T	phosphorylation	Y44[b]	1	1
37	50	L.SPEHVEVYERTHGQ.S			1	2
37	52	L.SPEHVEVYERTHGQSH.Y			1	
38	44	S.PEHVEVY.E			3	
49	54	H.GQSHYR.R			1	
P2 Isoform 2 (P04554-2)						
118	126	T.KLPGPLTPS.W	phosphorylation	T124[b]		1
118	127	T.KLPGPLTPSW.K			1	

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

Table 2. continued

[a]Abbreviations: P2 = protamine 2, S = serine residue, T = threonine residue, Y = tyrosine. residue. [b]Novel identified modified residues.

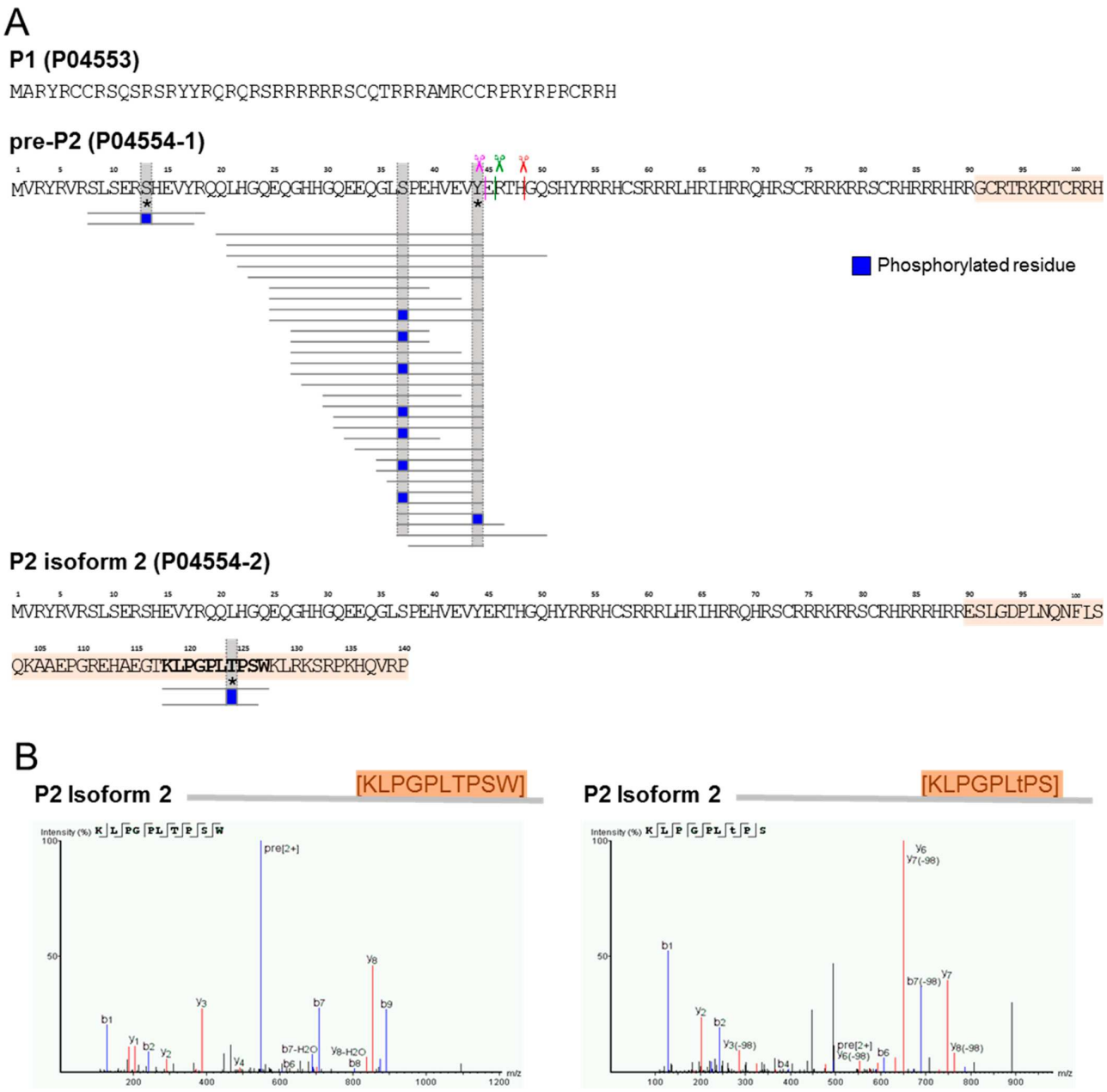


Figure 4. Protamine peptides identified by the bottom-up mass spectrometry (MS) approach. (A) No peptides were identified in the P1 amino acid sequence (upper box). The lower box shows part of the peptides located in the amino acid sequence of pre-P2 and P2 isoform 2, including their associated phosphorylated residues, represented with blue squares. In light orange is highlighted the difference between the peptide sequence of P2 and P2 isoform 2. Scissors indicate the proteolytic site for mature forms (HP2 in green, HP3 in red, and HP4 in pink). Asterisks mark novel modified residues. (B) P2 isoform 2 spectra corresponding to the two unique peptides identified.

employed in an attempt to reassign nonlocalized modifications. Figure 3B provides the best hits for site localization of each P1 and P2 family proteoforms. The corresponding fragmentation maps supporting the overall proteoform characterization are reported in Supporting Table 2 and Supporting Figures 4 and 5. However, in some cases, the detected fragment ions did not allow an accurate site localization.

A phosphorylated pattern was established for P1 through the top-down MS approach. In particular, MS1 identified, in decrease count number, mono-, di-, and triphosphorylated P1 proteoforms, in addition to unmodified P1 (Table 1, Figure 2, and Figure 3). Following this pipeline, MS2 analysis was able to confirm for P1 the presence of monophosphorylations at S11 or S13 and diphosphorylations at S11−S22 or S11−S29

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

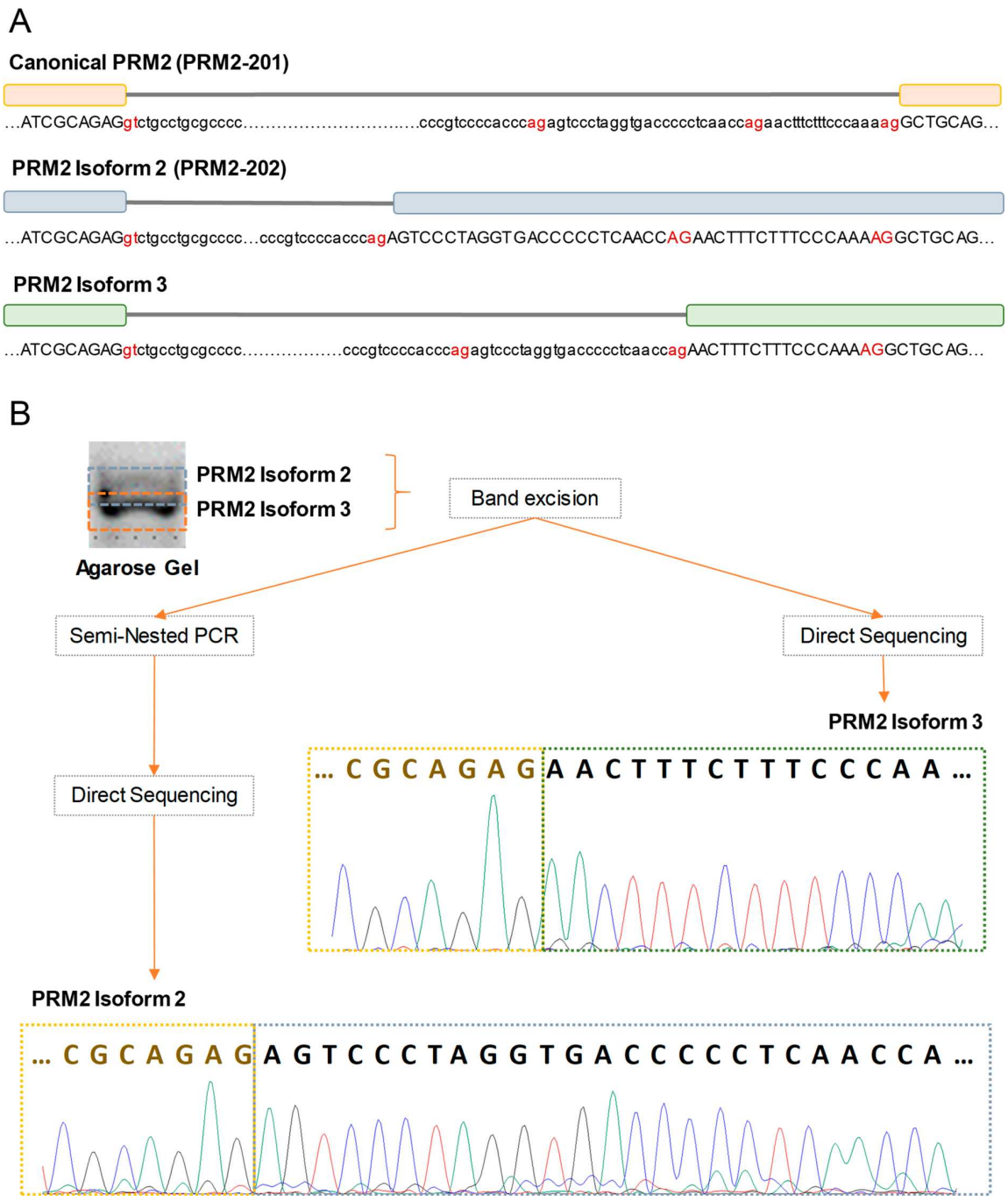


Figure 5. P2 isoform 2 (ENST00000435245.2) validated at the RNA level by direct sequencing. (A) Graphical representation of the two exons comprising *PRM2* gene with the two different transcripts derived from splice variants. Upper and lower cases correspond to exon and intron regions, respectively, and red letters indicate the splice sites. (B) P2 isoforms 2 and 3 separated by a 3% agarose gel were excised, and while direct sequencing directly sequenced P2 isoform 3, seminested PCR followed by direct sequencing was necessary to obtain the P2 isoform 2 sequence. The yellow box and brown letters indicate the final part of exon 1 common in all transcripts, and the P2 isoform 3 and isoform 2 splice sites are shown in the green and gray boxes, respectively.

(Supporting Table 2, Figure 3B, and Supporting Figure 4). Moreover, we identified a monophosphorylation of S11 in the truncated P1 proteoform (Figure 3B and Supporting Figure 6A). Additionally, pre-P2 was detected in its unmodified form as well as containing several PTMs. Specifically, MS2 fragmentation showed a phosphorylation pattern for pre-P2, including monophosphorylation, diphosphorylations, triphosphorylations, and tetraphosphorylations. However, proteoform-specific modified residues were not localized (Supporting Table 2, Figure 3B, and Supporting Figure 5). For P2 mature proteoforms, S59 was detected as phosphorylated for both HP2 and HP3 components, while no PTMs were identified in HP4 (Supporting Table 2, Figure 3B, and Supporting Figure 5). Total counts confirmed that unmodified HP2 was the most

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

Table 3. log_2(intensities) for P1, P2, and pre-P2 and the P1/P2 and pre-P2/P2 Ratios Compared to the Conventionally Acquired P1/P2 Ratio

	log_2(intensity)					
	P1 proteoforms	P2 mature proteoforms	pre-P2 proteoforms	P1/P2 ratio	pre-P2/P2 ratio	acid−urea PAGE P1/P2 ratio
b1	409.27	250.74	110.58	1.63	0.44	0.95
b2	325.38	187.39	384.20	1.74	2.05	0.89
b3	459.17	224.58	452.14	2.04	2.01	1.22
b4	374.71	222.28	427.27	1.69	1.92	0.92
b5	424.88	248.03	206.12	1.71	0.83	1.21
av	398.68	226.60	316.06	1.76	1.45	1.04
SD	51.00	25.50	149.87	0.16	0.76	0.16
CV (%)	12.80	11.30	47.40	9.10	52.40	15.40

identified mature P2 proteoform, followed by unmodified HP3, unmodified HP4, monophosphorylated HP2, and monophosphorylated HP3. From the truncated P2 proteoforms, only HPS1 was detected harboring modified residues (Table 1 and Figure 3). In particular, HPS1 proteoforms included either a monophosphorylation at S59 or the modification of the N-terminus Q34 to pyroglutamic acid (Supporting Table 2, Figure 3B, and Supporting Figure 5). The modification of the N-terminus glutamine residue to pyroglutamic acid is a reaction commonly found in plants and animals, including humans, and it has been proposed to be involved in the stabilization of proteins and peptides structures.[77] Complementary to these results, the bottom-up MS approach identified modifications in the sequence of pre-P2. Specifically, monophosphorylated peptides at S11, S37, or Y44 residues were detected. Additionally, P2 isoform 2 was found to be phosphorylated at the T124 residue (Table 2 and Figure 4).

Our results are supported by previous publications describing retained phosphorylated residues in protamines from mature sperm through different approaches, such as phosphoserine conversion followed by protein sequencing, electrospray mass spectrometry, or by similarity from other species. For instance, P1 has been described by others as phosphorylated at S9, S11, and S29 residues and pre-P2 as phosphorylated at S8, S10, S37, T47, S51, S59, and S73.[50−52,78] A phosphorylation pattern for P1, including mono-, di-, and triphosphorylation, as well as the presence of methylations and acetylations in HP3, was previously described through a top-down MS approach in human sperm, although no specific localization of the modified residues was provided.[52] In mice, Brunner et al. described different combinations of protamine PTMs using a combination of top-down and bottom-up MS approaches. Specifically, they identified phosphorylated, methylated, and acetylated residues for both protamines and confirmed their site location by a peptide-based approach.[57] Their results using the top-down method also revealed that multiple PTMs co-occur on individual protamines. In this respect, we have found different combinations of PTMs on the intact molecules, and only the most probable protamine proteoforms after performing a data-clustering correction and a further validation with Prosight Lite were annotated. In contrast to Brunner's work, our results highlight that phosphorylated proteoforms are predominant in the protamine family, even without performing an acetylation, methylation, or phosphorylation sample enrichment with strong cation exchange chromatography.

The conventional analysis of protamines in infertile patients includes the determination of relative protein quantities by acid−urea PAGE and generation of P1/P2 ratios. However, this procedure is unable to detect protamine proteoforms, despite some improvements using two-dimensional gel electrophoresis.[79] Therefore, the proteomic strategies described here are revealed as the most suitable to detect alterations in the protamine proteoforms profile associated with male infertility. Recently, the crucial role of P2 dephosphorylation in male fertility has been described in a mouse model,[49] pointing out the existence of a tightly regulated balance of phosphorylation/dephosphorylation state during sperm maturation. Impairments in this dynamic network could then be suggested as the underlying cause of different types of infertile patients, which demands future protamine proteoforms comparative studies in order to elucidate its role. Likewise, our results showed that unmodified proteoforms are more abundant than phosphorylated proteoforms, supporting the hypothesis that a high dephosphorylation wave occurs during sperm maturation.[1,8,46] Remaining phosphorylated residues in mature human sperm could either be simple spermatogenesis remnants randomly found across some protamine molecules or, in contrast, act as specific epigenetic marks during early embryo development. Since the attachment of phosphorylated protamines to the DNA is not as strong as for unmodified protamines, due to the reduction of positive charge, phosphorylated protamines could mark the preferential paternal protamine replacement by maternal histones during early embryo development.[7] However, the function of these retained phosphorylated protamines beyond fertilization, if any, remains to be elucidated.

Surprisingly, through the top-down MS approach a +61 Da modification was widely identified solely or in combination with other PTMs across P1, HP2, HP3, HPS1, and HPS2 (Table 1, Figure 3, and Supporting Figure 3). In particular, MS2 fragmentation showed the localization of this modification in residues C31 and/or C51 of P1, as monomodified or in conjunction with mono-, di-, and triphosphorylation, and in residues C82 of HP2 and HPS2, C74 of HP3, and C92 of HPS1 (Supporting Table 2, Figure 3B, and Supporting Figures 4 and 5). In HPS1, this modification was detected either solely or in combination with the Q34 modification to pyroglutamic acid (Figure 3B). The incorporation of 61 Da to the molecular weight of protamines might correspond to the presence of a Zn^{2+} ion attached to protamine molecules. It has been demonstrated that low levels of Zn^{2+} in seminal plasma negatively correlate with sperm quality and fertility.[80,81] Moreover, it has been also proposed that Zn^{2+} could aid the stabilization of sperm chromatin.[7] The attachment of Zn^{2+} to P2 has been previously demonstrated,[82] and a stoichiometric binding of one P2 molecule to one Zn^{2+} ion has been

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

proposed.[30] In addition, a zinc-finger motif in P2 (Cys2-His2) has been described that could allow P2 proteins to act as zinc-finger proteins.[83] In the present study, we detected the +61 Da modification in P1, although the attachment of Zn^{2+} to P1 has not been previously described. However, further studies based on inductively coupled plasma (ICP) masses are required in order to clarify whether the presence of this modification in both protamines corresponds to Zn^{2+} ion, as well as its potentially related function.

Determination of the P1/P2 Ratio and the Pre-P2/P2 Ratio Using Top-Down MS Data

The conventional P1/P2 ratio obtained by acid−urea PAGE-based experiments has been widely evaluated to assess the fertility state in andrology research. However, although it has been linked to sperm penetration score, fertilization rates, and pregnancy outcome,[24,25] controversial results are found.[26] Conventional techniques based on acid−urea PAGE show limitations, such as potential differences in the staining affinities of the dyes, as well as the difficulty in discerning between proteoforms that differ only by small mass or isoelectric point changes. These limitations could be overcome with top-down MS-derived data, including detailed proteoform information. By using LC−MS intensities of P1 proteoforms and mature P2 proteoforms, an average P1/P2 ratio of 1.76 with a standard deviation of 0.16 was revealed (Table 3). This ratio is notably higher than the one calculated using the conventional acid−urea PAGE P1/P2 ratio, for which an average of 1.04 with a standard deviation of 0.16 was obtained (Table 3). One potential explanation for this high MS-derived ratio may be the better chromatographic efficiency and ionization response of P1 proteoforms in comparison to P2 proteoforms, related to their particular physical−chemical properties. Despite the technical limitations associated with both techniques, P1/P2 ratios derived by MS and acid−urea PAGE displayed similar values among normoozospermic individuals, showing low standard deviation values across the samples assessed in this study (Table 3). Taking into account the valuable proteoform-specific information obtained when applying the top-down based approach, the evaluation of potential dysregulations in the MS-derived P1/P2 ratio in different types of infertile patients, as previously done with conventional approaches (reviewed in refs 24−26), is demanded.

Alterations in pre-P2 have also been reported in infertile patients using conventional approaches.[23,28,29,31,84,85] Specifically, an association between a high pre-P2/P2 mature proteoforms ratio (pre-P2/P2) with alterations in seminal parameters, specifically low sperm counts, low sperm motility, and abnormal morphology, was established.[28] Interestingly, when assisted reproductive outcomes were evaluated, a low pre-P2/P2 ratio was linked to poor pregnancy success.[29] These results suggest that the deregulation of specific immature proteoforms not detected by standard methods could be the underlying cause of pre-P2/P2 alterations. For that reason, the ratio between the intensities of pre-P2 proteoforms and P2 mature proteoforms (pre-P2/P2) was also calculated here from our top-down MS-derived data. The obtained MS-derived pre-P2/P2 ratios ranged from 0.44 to 2.05, being on average 1.45 (Table 3). In contrast to the clear similarity of the MS-derived P1/P2 ratio among our samples, the MS-derived pre-P2/P2 ratio differs among individuals, being established as two different subgroups (Table 3). The first subgroup included 3 out the 5 samples showing a pre-P2/P2 ratio of approximately 2, while the second subgroup was form by the remaining two samples, showing a lower pre-P2/P2 ratio of around 0.6, due to lower intensities of pre-P2 proteoforms. These results suggest that normozoospermic infertile patients could be stratified on the basis of the pre-P2/P2 ratio. However, it remains necessary to elucidate whether a dysregulation of the MS-derived pre-P2/P2 ratio might be associated with assisted reproductive outcomes. Altogether this suggests the accuracy of the top-down proteoform profile as an advantage to allow on to determine alterations in specific proteoforms potentially associated with different phenotypes of infertile men.

CONCLUSIONS

In this study, we demonstrated that the combination of top-down and bottom-up MS approaches is suitable to study the sperm-specific protamine proteoform profile, giving complementary information with low interindividual variability among the analyzed patients. These approaches have permitted, for the first time, establishment of the normal phosphorylation pattern for human protamines, by identifying known and novel phosphorylated residues for P1, pre-P2, HP2, and HP3. Remarkably, our top-down MS results suggest that the normal protamine proteoform profile comprises unmodified P1 as the most widely identified proteoform, followed by unmodified mature HP2, monophosphorylated P1, unmodified HP3, diphosphorylated P1, unmodified HP4, monophosphorylated HP2, and monophosphorylated HP3. Moreover, previously detected and novel truncated proteoforms have been identified for both P1 and pre-P2, some of which containg modified residues, possibly generated by post-translational proteolytic cleavage. Interestingly, in addition to phosphorylation, both protamines harbor a +61 Da modification that might correspond to Zn^{2+} ion, although further specific analyses are required to clarify this association. Furthermore, the proteinase-K-based bottom-up MS approach unequivocally identified the nonannotated P2 isoform 2. Since the utility of the conventional P1/P2 ratio based on their relative quantity is limited and it cannot identify alterations in specific protamine proteoforms, our results increase the current knowledge of human protamines by setting up the normal profile of protamine proteoforms in human individual sperm samples with normal semen parameters. Furthermore, we evaluated P1/P2 and pre-P2/P2 ratios using top-down MS-derived data. In contrast to the data obtained by using PAGE-based methods, these ratios would aid to determine specific proteoform alterations potentially related to different infertility phenotypes. The study of proteoform dysregulations in different types of patients, particularly phosphorylation/dephosphorylation dynamics as well as alterations in P1/P2 and MS-derived pre-P2/P2 ratios, opens a window to determine their potential role in male fertility, along with the potential function of protamine modifications as epigenetic marks beyond fertilization.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.9b00499.

Supporting Table 1 summarizes the primer sets used to verify P2 isoform 2 at the RNA level; Supporting Table 2 details the manual MS2 fragmentation of the proteo-

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

forms shown in Supporting Figures 4 and 5 for P1 and P2; Supporting Figure 1 shows the pipeline applied for the top-down MS and the bottom-up MS approaches; Supporting Figure 2 schematizes the overall strategy employed to identify PRM2 isoform 2 at the RNA level; Supporting Figure 3 shows with a scatter plot each biological replicate proteoform profile; Supporting Figures 4 and 5 show the fragmentation of the identified proteoforms for P1 and P2, respectively; Supporting Figure 6 displays spectra corresponding to the truncated forms of P1 and P2; Supporting Figure 7 illustrates "CLUSTAL O" alignment in different species of P2 isoform 2; and Supporting Figure 8 shows specific modified residues for P1 and P2 employing two different automatized engines (PDF)

AUTHOR INFORMATION

Corresponding Author

*Telephone: +34 93 402 18 77. Fax: +34 93 403 52 78. E-mail: roliva@ub.edu.

ORCID

Rafael Oliva: 0000-0003-4876-2410

Author Contributions

⊥A.S.-V., M.G., and M.J. have contributed equally to this work. A.S.-V., M.J., J.C., and RO participated in the study design, sample preparation, data analysis, and manuscript preparation. J.L.B. participated in sample collection and manuscript preparation. M.G., G.A.-G., M.V., L.V., and M.V. participated in LC−MS/MS, data analysis, and manuscript preparation. The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

The mass spectrometry data have been deposited to the PRIDE Archive (http://www.ebi.ac.uk/pride/archive/) via the PRIDE partner repository with the data set identifier PXD014618.

ACKNOWLEDGMENTS

This work was supported by grants from EUGIN-UB Research Excellence Program (EU-REP 2014), the Spanish Ministry of Economy and Competitiveness (Ministerio de Economia y Competitividad; Fondos FEDER 'Una manera de hacer Europa', PI 13/00699, and PI 16/00346), and Fundación Salud 2000 (SERONO 13-015) to R.O. M.J. is supported by a Postdoctoral Fellowship from the Government of Catalonia (Generalitat de Catalunya, pla estratègic de recerca i innovació en salut, PERIS 2016−2020, SLT SLT002/16/00337). J.C. is supported by the Sara Borrell Postdoctoral Fellowship from the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad, Acción Estratégica en Salud, CD17/00109). The Mass Spectrometry and Proteomics Core Facility where the proteomic experiments have been conducted is a ProteoRed lab member supported by PT17/0019/-ISCIII-SGEFI/ERDF. The authors recognize Raquel Ferreti and Alicia Diez for their assistance in the routine seminograms and sample collection.

ABBREVIATIONS

PRM1, protamine 1 gene; PRM2, protamine 2 gene; TNP2, transition nuclear protein 2; P1, protamine 1 protein; P2, protamine 2 protein; pre-P2, protamine 2 precursor; HP2, protamine 2 mature form 2; HP3, protamine 2 mature form 3; HP4, protamine 2 mature form 4; PAGE, polyacrylamide gel electrophoresis; PTM, post-translational modification; SRPK1, phosphorylated by SRSF protein kinase 1; CAMK4, calcium/calmodulin dependent protein kinase IV; Hspa4l, heat shock protein family A (Hsp70) member 4 like; Ppp1cc2, protein phosphatase 1 catalytic subunit gamma isoform 2; HCl, chloridric acid; $MgCl_2$, magnesium chloride; NH_4HCO_3, ammonium bicarbonate; IAA, iodoacetamide; $CaCl_2$, calcium chloride; ACN, acetonitrile; FA, formic acid; DDA, data-dependent acquisition; HCD, high-collision dissociation; LC−MS/MS, liquid chromatography coupled to tandem mass spectrometry; FDR, false discovery rate

REFERENCES

(1) Oliva, R. Protamines and Male Infertility. *Hum. Reprod. Update* **2006**, *12* (4), 417−435.

(2) Kimmins, S.; Sassone-Corsi, P. Chromatin Remodelling and Epigenetic Features of Germ Cells. *Nature* **2005**, *434* (7033), 583−589.

(3) Balhorn, R. The Protamine Family of Sperm Nuclear Proteins. *Genome Biol.* **2007**, *8* (9), 227.

(4) Bao, J.; Bedford, M. T. Epigenetic Regulation of the Histone-to-Protamine Transition during Spermiogenesis. *Reproduction* **2016**, *151* (5), R55−R70.

(5) Barral, S.; Morozumi, Y.; Tanaka, H.; Montellier, E.; Govin, J.; de Dieuleveult, M.; Charbonnier, G.; Couté, Y.; Puthier, D.; Buchou, T.; et al. Histone Variant H2A.L.2 Guides Transition Protein-Dependent Protamine Assembly in Male Germ Cells. *Mol. Cell* **2017**, *66* (1), 89−101.e8.

(6) Balhorn, R.; Corzett, M.; Mazrimas, J. A. Formation of Intraprotamine Disulfides in Vitro. *Arch. Biochem. Biophys.* **1992**, *296* (2), 384−393.

(7) Björndahl, L.; Kvist, U. Human Sperm Chromatin Stabilization: A Proposed Model Including Zinc Bridges. *Mol. Hum. Reprod.* **2010**, *16* (1), 23−29.

(8) Oliva, R.; Dixon, G. H. Vertebrate Protamine Genes and the Histone-to-Protamine Replacement Reaction. *Prog. Nucleic Acid Res. Mol. Biol.* **1991**, *40* (C), 25−94.

(9) Oliva, R.; Castillo, J. Proteomics and the Genetics of Sperm Chromatin Condensation. *Asian J. Androl.* **2011**, *13* (1), 24−30.

(10) Oliva, R.; Bazett-Jones, D.; Mezquita, C.; Dixon, G. H. Factors Affecting Nucleosome Disassembly by Protamines in Vitro. Histone Hyperacetylation and Chromatin Structure, Time Dependence, and the Size of the Sperm Nuclear Proteins. *J. Biol. Chem.* **1987**, *262* (35), 17016−17025.

(11) Oliva, R.; Mezquita, C. Marked Differences in the Ability of Distinct Protamines To Disassemble Nucleosomal Core Particles in Vitro. *Biochemistry* **1986**, *25* (21), 6508−6511.

(12) Kasinsky, H. E.; Eirín-López, J. M.; Ausió, J. Protamines: Structural Complexity, Evolution and Chromatin Patterning. *Protein Pept. Lett.* **2011**, *18* (8), 755−771.

(13) Domenjoud, L.; Kremling, H.; Burfeind, P.; Maier, W. M.; Engel, W. On the Expression of Protamine Genes in the Testis of Man and Other Mammals. *Andrologia* **1991**, *23* (5), 333−337.

(14) Nelson, J. E.; Krawetz, S. A. Mapping the Clonally Unstable Recombinogenic PRM1→PRM2→TNP2 Region of Human 16p 13.2. *DNA Sequence* **1995**, *5* (3), 163−168.

(15) McKay, D. J.; Renaux, B. S.; Dixon, G. H. The Amino Acid Sequence of Human Sperm Protamine P1. *Biosci. Rep.* **1985**, *5* (5), 383−391.

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

(16) McKay, D. J.; Renaux, B. S.; Dixon, G. H. Human Sperm Protamines: Amino-acid Sequences of Two Forms of Protamine P2. *Eur. J. Biochem.* **1986**, *156* (1), 5−8.
(17) De Yebra, L.; Ballesca, J. L.; Vanrell, J. A.; Bassas, L.; Oliva, R. Complete Selective Absence of Protamine P2 in Humans. *J. Biol. Chem.* **1993**, *268* (14), 10553−10557.
(18) Yelick, P.; Balhorn, R.; Johnson, P.; Corzett, M.; Mazrimas, J.; Kleene, K.; Hecht, N. Mouse Protamine 2 Is Synthesized as a Precursor Whereas Mouse Protamine 1 Is Not. *Mol. Cell. Biol.* **1987**, *7* (6), 2173.
(19) Smith, L. M.; Kelleher, N. L. Consortium for Top Down Proteomics. Proteoform: A Single Term Describing Protein Complexity. *Nat. Methods* **2013**, *10* (3), 186−187.
(20) Jodar, M.; Oliva, R. Protamine Alterations in Human Spermatozoa. *Adv. Exp. Med. Biol.* **2014**, *791*, 83−102.
(21) Arkhis, A.; Martinage, A.; Sautiere, P.; Chevaillier, P. Molecular Structure of Human Protamine P4 (HP4), a Minor Basic Protein of Human Sperm Nuclei. *Eur. J. Biochem.* **1991**, *200* (2), 387−392.
(22) Balhorn, R.; Corzett, M.; Mazrimas, J.; Stanker, L. H.; Wyrobek, A. High-Performance Liquid Chromatographic Separation and Partial Characterization of Human Protamines 1, 2, and 3. *Biotechnol. Appl. Biochem.* **1987**, *9* (1), 82−88.
(23) de Mateo, S.; Ramos, L.; de Boer, P.; Meistrich, M.; Oliva, R. Protamine 2 Precursors and Processing. *Protein Pept. Lett.* **2011**, *18* (8), 778−785.
(24) Barrachina, F.; Soler-Ventura, A.; Oliva, R.; Jodar, M. Sperm Nucleoproteins (Histones and Protamines). In *A Clinician's Guide to Sperm DNA and Chromatin Damage*; Springer International Publishing: Cham, Switzerland, 2018; pp 31−51.
(25) Soler-Ventura, A.; Castillo, J.; de la Iglesia, A.; Jodar, M.; Barrachina, F.; Ballesca, J. L.; Oliva, R. Mammalian Sperm Protamine Extraction and Analysis: A Step-By-Step Detailed Protocol and Brief Review of Protamine Alterations. *Protein Pept. Lett.* **2018**, *25* (5), 424−433.
(26) Ni, K.; Spiess, A.-N.; Schuppe, H.-C.; Steger, K. The Impact of Sperm Protamine Deficiency and Sperm DNA Damage on Human Male Fertility: A Systematic Review and Meta-Analysis. *Andrology* **2016**, *4* (5), 789−799.
(27) Nanassy, L.; Liu, L.; Griffin, J.; Carrell, D. T. The Clinical Utility of the Protamine 1/protamine 2 Ratio in Sperm. *Protein Pept. Lett.* **2011**, *18* (8), 772−777.
(28) Torregrosa, N.; Domínguez-Fandos, D.; Camejo, M. I.; Shirley, C. R.; Meistrich, M. L.; Ballescà, J. L.; Oliva, R. Protamine 2 Precursors, Protamine 1/protamine 2 Ratio, DNA Integrity and Other Sperm Parameters in Infertile Patients. *Hum. Reprod.* **2006**, *21* (8), 2084−2089.
(29) de Mateo, S.; Gázquez, C.; Guimerà, M.; Balasch, J.; Meistrich, M. L.; Ballescà, J. L.; Oliva, R. Protamine 2 Precursors (Pre-P2), Protamine 1 to Protamine 2 Ratio (P1/P2), and Assisted Reproduction Outcome. *Fertil. Steril.* **2009**, *91* (3), 715−722.
(30) Bench, G.; Corzett, M. H.; Kramer, C. E.; Grant, P. G.; Balhorn, R. Zinc Is Sufficiently Abundant within Mammalian Sperm Nuclei to Bind Stoichiometrically with Protamine 2. *Mol. Reprod. Dev.* **2000**, *56* (4), 512−519.
(31) De Yebra, L.; Ballescá, J. L.; Vanrell, J. A.; Corzett, M.; Balhorn, R.; Oliva, R. Detection of P2 Precursors in the Sperm Cells of Infertile Patients Who Have Reduced Protamine P2 Levels. *Fertil. Steril.* **1998**, *69* (4), 755−759.
(32) Castillo, J.; Jodar, M.; Oliva, R. The Contribution of Human Sperm Proteins to the Development and Epigenome of the Preimplantation Embryo. *Hum. Reprod. Update* **2018**, *24* (5), 535−555.
(33) Siklenka, K.; Erkek, S.; Godmann, M.; Lambrot, R.; McGraw, S.; Lafleur, C.; Cohen, T.; Xia, J.; Suderman, M.; Hallett, M. Disruption of Histone Methylation in Developing Sperm Impairs Offspring Health Transgenerationally. *Science* **2015**, *350* (6261), aab2006.
(34) Carrell, D. T.; Emery, B. R.; Hammoud, S. Altered Protamine Expression and Diminished Spermatogenesis: What Is the Link? *Hum. Reprod. Update* **2007**, *13* (3), 313−327.
(35) Hammoud, S. S.; Nix, D. A.; Zhang, H.; Purwar, J.; Carrell, D. T.; Cairns, B. R. Distinctive Chromatin in Human Sperm Packages Genes for Embryo Development. *Nature* **2009**, *460* (7254), 473−478.
(36) Castillo, J.; Amaral, A.; Azpiazu, R.; Vavouri, T.; Estanyol, J. M.; Ballesca, J. L.; Oliva, R. Genomic and Proteomic Dissection and Characterization of the Human Sperm Chromatin. *Mol. Hum. Reprod.* **2014**, *20* (11), 1041−1053.
(37) Arpanahi, A.; Brinkworth, M.; Iles, D.; Krawetz, S. A.; Paradowska, A.; Platts, A. E.; Saida, M.; Steger, K.; Tedder, P.; Miller, D. Endonuclease-Sensitive Regions of Human Spermatozoal Chromatin Are Highly Enriched in Promoter and CTCF Binding Sequences. *Genome Res.* **2009**, *19* (8), 1338−1349.
(38) Royo, H.; Stadler, M. B.; Peters, A. H. F. M. Alternative Computational Analysis Shows No Evidence for Nucleosome Enrichment at Repetitive Sequences in Mammalian Spermatozoa. *Dev. Cell* **2016**, *37* (1), 98−104.
(39) Samans, B.; Yang, Y.; Krebs, S.; Sarode, G. V.; Blum, H.; Reichenbach, M.; Wolf, E.; Steger, K.; Dansranjavin, T.; Schagdarsurengin, U. Uniformity of Nucleosome Preservation Pattern in Mammalian Sperm and Its Connection to Repetitive DNA Elements. *Dev. Cell* **2014**, *30* (1), 23−35.
(40) Krejčí, J.; Stixová, L.; Pagáčová, E.; Legartová, S.; Kozubek, S.; Lochmanová, G.; Zdráhal, Z.; Sehnalová, P.; Dabravolski, S.; Hejátko, J.; et al. Post-Translational Modifications of Histones in Human Sperm. *J. Cell. Biochem.* **2015**, *116* (10), 2195−2209.
(41) Baker, M. A. Proteomics of Post-Translational Modifications of Mammalian Spermatozoa. *Cell Tissue Res.* **2016**, *363* (1), 279−287.
(42) Papoutsopoulou, S.; Nikolakaki, E.; Chalepakis, G.; Kruft, V.; Chevaillier, P.; Giannakouros, T. SR Protein-Specific Kinase 1 Is Highly Expressed in Testis and Phosphorylates Protamine 1. *Nucleic Acids Res.* **1999**, *27* (14), 2972−2980.
(43) Green, G. R.; Balhorn, R.; Poccia, D. L.; Hecht, N. B. Synthesis and Processing of Mammalian Protamines and Transition Proteins. *Mol. Reprod. Dev.* **1994**, *37* (3), 255−263.
(44) Pirhonen, A.; Linnala-Kankkunen, A.; Mäenpää, P. H. P2 Protamines Are Phosphorylated in Vitro by Protein Kinase C, Whereas P1 Protamines Prefer cAMP-Dependent Protein Kinase. A Comparative Study of Five Mammalian Species. *Eur. J. Biochem.* **1994**, *223* (1), 165−169.
(45) Wu, J. Y.; Ribar, T. J.; Cummings, D. E.; Burton, K. A.; McKnight, G. S.; Means, A. R. Spermiogenesis and Exchange of Basic Nuclear Proteins Are Impaired in Male Germ Cells Lacking Camk4. *Nat. Genet.* **2000**, *25* (4), 448−452.
(46) Carrell, D. T.; Emery, B. R.; Hammoud, S. Altered Protamine Expression and Diminished Spermatogenesis: What Is the Link? *Hum. Reprod. Update* **2007**, *13* (3), 313−327.
(47) Wu, J. Y.; Ribar, T. J.; Cummings, D. E.; Burton, K. A.; McKnight, G. S.; Means, A. R. Spermiogenesis and Exchange of Basic Nuclear Proteins Are Impaired in Male Germ Cells Lacking Camk4. *Nat. Genet.* **2000**, *25* (4), 448−452.
(48) Gusse, M.; Sautiere, P.; Belaiche, D.; Martinage, A.; Roux, C.; Dadoune, J. P.; Chevaillier, P. Purification and Characterization of Nuclear Basic Proteins of Human Sperm. *Biochim. Biophys. Acta, Gen. Subj.* **1986**, *884* (1), 124−134.
(49) Itoh, K.; Kondoh, G.; Miyachi, H.; Sugai, M.; Kaneko, Y.; Kitano, S.; Watanabe, H.; Maeda, R.; Imura, A.; Liu, Y.; et al. Dephosphorylation of Protamine 2 at Serine 56 Is Crucial for Murine Sperm Maturation in Vivo. *Sci. Signaling* **2019**, *12* (574), eaao7232.
(50) Pirhonen, A.; Linnala-Kankkunen, A.; Mäenpää, P. H. Identification of Phosphoseryl Residues in Protamines from Mature Mammalian Spermatozoa1. *Biol. Reprod.* **1994**, *50* (5), 981−986.
(51) Chira, F.; Arkhis, A.; Martinage, A.; Jaquinod, M.; Chevaillier, P.; Sautière, P. Phosphorylation of Human Sperm Protamines HP1 and HP2: Identification of Phosphorylation Sites. *Biochim. Biophys. Acta* **1993**, *1203* (1), 109−114.

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

(52) Castillo, J.; Estanyol, J. M.; Ballescà, J.; Oliva, R. Human Sperm Chromatin Epigenetic Potential: Genomics, Proteomics, and Male Infertility. *Asian J. Androl.* **2015**, *17* (4), 601.
(53) Amaral, A.; Castillo, J.; Ramalho-Santos, J.; Oliva, R. The Combined Human Sperm Proteome: Cellular Pathways and Implications for Basic and Clinical Science. *Hum. Reprod. Update* **2014**, *20* (1), 40−62.
(54) Smith, L. M.; Kelleher, N. L. Proteoforms as the next Proteomics Currency. *Science* **2018**, *359* (6380), 1106−1107.
(55) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; et al. How Many Human Proteoforms Are There? *Nat. Chem. Biol.* **2018**, *14* (3), 206−214.
(56) Bonet-Costa, C.; Vilaseca, M.; Diema, C.; Vujatovic, O.; Vaquero, A.; Omeñaca, N.; Castejón, L.; Bernués, J.; Giralt, E.; Azorín, F. Combined Bottom-up and Top-down Mass Spectrometry Analyses of the Pattern of Post-Translational Modifications of Drosophila Melanogaster Linker Histone H1. *J. Proteomics* **2012**, *75* (13), 4124−4138.
(57) Brunner, A. M.; Nanni, P.; Mansuy, I. M. Epigenetic Marking of Sperm by Post-Translational Modification of Histones and Protamines. *Epigenetics Chromatin* **2014**, *7* (1), 2.
(58) Cooper, T. G. *WHO Laboratory Manual for the Examination and Processing of Human Semen*, 5th ed.; WHO, 2010.
(59) Mengual, L.; Ballescà, J. L.; Ascaso, C.; Oliva, R. Marked Differences in Protamine Content and P1/P2 Ratios. *J. Androl.* **2003**, *24* (3), 438−447.
(60) Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *KDD-96 Proceedings*; AAAI, 1996.
(61) Zabrouskov, V.; Senko, M. W.; Du, Y.; Leduc, R. D.; Kelleher, N. L. New and Automated MSn Approaches for Top-down Identification of Modified Proteins. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 2027−2038.
(62) Liu, X.; Inbar, Y.; Dorrestein, P. C.; Wynne, C.; Edwards, N.; Souda, P.; Whitelegge, J. P.; Bafna, V.; Pevzner, P. A. Deconvolution and Database Search of Complex Tandem Mass Spectra of Intact Proteins: A Combinatorial Approach. *Mol. Cell. Proteomics* **2010**, *9* (12), 2772−2782.
(63) LeDuc, R. D.; Taylor, G. K.; Kim, Y. B.; Januszyk, T. E.; Bynum, L. H.; Sola, J. V.; Garavelli, J. S.; Kelleher, N. L. ProSight PTM: An Integrated Environment for Protein Identification and Characterization by Top-down Mass Spectrometry. *Nucleic Acids Res.* **2004**, *32*, W340−W345.
(64) Kou, Q.; Xun, L.; Liu, X. TopPIC: A Software Tool for Top-down Mass Spectrometry-Based Proteoform Identification and Characterization. *Bioinformatics* **2016**, *32* (22), 3495−3497.
(65) Fellers, R. T.; Greer, J. B.; Early, B. P.; Yu, X.; LeDuc, R. D.; Kelleher, N. L.; Thomas, P. M. ProSight Lite: Graphical Software to Analyze Top-down Mass Spectrometry Data. *Proteomics* **2015**, *15* (7), 1235−1238.
(66) Goodrich, R. J.; Anton, E.; Krawetz, S. A. Isolating mRNA and Small Noncoding RNAs from Human Sperm. *Methods Mol. Biol.* **2013**, *927*, 385−396.
(67) Jodar, M.; Sendler, E.; Moskovtsev, S. I.; Librach, C. L.; Goodrich, R.; Swanson, S.; Hauser, R.; Diamond, M. P.; Krawetz, S. A. Absence of Sperm RNA Elements Correlates with Idiopathic Male Infertility. *Sci. Transl. Med.* **2015**, *7* (295), 295re6.
(68) Sautière, P.; Martinage, A.; Bélaiche, D.; Arkhis, A.; Chevaillier, P. Comparison of the Amino Acid Sequences of Human Protamines HP2 and HP3 and of Intermediate Basic Nuclear Proteins HPS1 and HPS2. Structural Evidence That HPS1 and HPS2 Are pro-Protamines. *J. Biol. Chem.* **1988**, *263* (23), 11059−11062.
(69) Alimi, E.; Martinage, A.; Arkhis, A.; Belaiche, D.; Sautiere, P.; Chevaillier, P. Amino Acid Sequence of the Human Intermediate Basic Protein 2 (HPI2) from Sperm Nuclei. Structural Relationship with Protamine P2. *Eur. J. Biochem.* **1993**, *214* (2), 445−450.
(70) Martinage, A.; Arkhis, A.; Alimi, E.; Sautière, P.; Chevaillier, P. Molecular Characterization of Nuclear Basic Protein HPI1, a Putative Precursor of Human Sperm Protamines HP2 and HP3. *Eur. J. Biochem.* **1990**, *191* (2), 449−451.
(71) Carré-Eusèbe, D.; Lederer, F.; Lê, K. H. D.; Elsevier, S. M. Processing of the Precursor of Protamine P2 in Mouse. Peptide Mapping and *N* -Terminal Sequence Analysis of Intermediates. *Biochem. J.* **1991**, *277* (1), 39−45.
(72) Chauvière, M.; Martinage, A.; Debarle, M.; Sautière, P.; Chevaillier, P. Molecular Characterization of Six Intermediate Proteins in the Processing of Mouse Protamine P2 Precursor. *Eur. J. Biochem.* **1992**, *204* (2), 759−765.
(73) The UniProt Consortium.. UniProt: A Worldwide Hub of Protein Knowledge. *Nucleic Acids Res.* **2019**, *47* (D1), D506−D515.
(74) Gaudet, P.; Michel, P.-A.; Zahn-Zabal, M.; Britan, A.; Cusin, I.; Domagalski, M.; Duek, P. D.; Gateau, A.; Gleizes, A.; Hinard, V.; et al. The neXtProt Knowledgebase on Human Proteins: 2017 Update. *Nucleic Acids Res.* **2017**, *45* (D1), D177−D182.
(75) Zerbino, D. R.; Achuthan, P.; Akanni, W.; Amode, M. R.; Barrell, D.; Bhai, J.; Billis, K.; Cummins, C.; Gall, A.; Girón, C. G.; et al. Ensembl 2018. *Nucleic Acids Res.* **2018**, *46* (D1), D754−D761.
(76) LeDuc, R. D.; Fellers, R. T.; Early, B. P.; Greer, J. B.; Thomas, P. M.; Kelleher, N. L. The C-Score: A Bayesian Framework to Sharply Improve Proteoform Scoring in High-Throughput Top down Proteomics. *J. Proteome Res.* **2014**, *13* (7), 3231−3240.
(77) Schilling, S.; Wasternack, C.; Demuth, H.-U. Glutaminyl Cyclases from Animals and Plants: A Case of Functionally Convergent Protein Evolution. *Biol. Chem.* **2008**, *389* (8), 983−991.
(78) Hornbeck, P. V.; Zhang, B.; Murray, B.; Kornhauser, J. M.; Latham, V.; Skrzypek, E. PhosphoSitePlus, 2014: Mutations, PTMs and Recalibrations. *Nucleic Acids Res.* **2015**, *43* (D1), D512−D520.
(79) Yoshii, T.; Kuji, N.; Komatsu, S.; Iwahashi, K.; Tanaka, Y.; Yoshida, H.; Wada, A.; Yoshimura, Y. Fine Resolution of Human Sperm Nucleoproteins by Two-Dimensional Electrophoresis. *Mol. Hum. Reprod.* **2005**, *11* (9), 677−681.
(80) Fallah, A.; Mohammad-Hasani, A.; Colagar, A. H. Zinc Is an Essential Element for Male Fertility: A Review of Zn Roles in Men's Health, Germination, Sperm Quality, and Fertilization. *J. Reprod. Infertil.* **2018**, *19* (2), 69−81.
(81) Colagar, A. H.; Marzony, E. T.; Chaichi, M. J. Zinc Levels in Seminal Plasma Are Associated with Sperm Quality in Fertile and Infertile Men. *Nutr. Res. (N. Y., NY, U. S.)* **2009**, *29* (2), 82−88.
(82) Bal, W.; Dyba, M.; Szewczuk, Z.; Jezowska-Bojczuk, M.; Lukszo, J.; Ramakrishna, G.; Kasprzak, K. S. Differential Zinc and DNA Binding by Partial Peptides of Human Protamine HP2. *Mol. Cell. Biochem.* **2001**, *222* (1−2), 97−106.
(83) Bianchi, F.; Rousseaux-Prevost, R.; Sautiere, P.; Rousseaux, J. P2 Protamines from Human Sperm Are Zinc -Finger Proteins with One Cys2His2Motif. *Biochem. Biophys. Res. Commun.* **1992**, *182* (2), 540−547.
(84) de Mateo, S.; Martínez-Heredia, J.; Estanyol, J. M.; Domíguez-Fandos, D.; Vidal-Taboada, J. M.; Ballescà, J. L.; Oliva, R. Marked Correlations in Protein Expression Identified by Proteomic Analysis of Human Spermatozoa. *Proteomics* **2007**, *7* (23), 4264−4277.
(85) Bench, G.; Corzett, M. H.; De Yebra, L.; Oliva, R.; Balhorn, R. Protein and DNA Contents in Sperm from an Infertile Human Male Possessing Protamine Defects That Vary over Time. *Mol. Reprod. Dev.* **1998**, *50* (3), 345−353.

DOI: 10.1021/acs.jproteome.9b00499
J. Proteome Res. 2020, 19, 221−237

3.4. Treball 4

Canvis proteòmics de l'espermatozoide associats a la qualitat embrionària primerenca després de ICSI

Reproductive Biomedicine Online

2020 May;40(5):700-710

PMID: 32444165

Sperm proteomic changes associated with early embryo quality after ICSI

Jodar M*, Attardo-Parrinello C*, Soler-Ventura A*, Barrachina F, Delgado-Dueñas D, Cívico S, Calafell JM, Ballescà JL, Oliva R.

* Aquests autors han contribuït per igual a aquest treball

Canvis proteòmics de l'espermatozoide associats a la qualitat embrionària primerenca després de ICSI

Objectiu: Identificar proteïnes de l'espermatozoide que podrien tenir un paper important en l'embriogènesi. Per dur-ho a terme, es va correlacionar l'abundància de les proteïnes de l'espermatozoide amb la qualitat dels embrions que deriven de fecundació *in vitro* amb injecció intracitoplasmàtica (FIV-ICSI), provinents d'individus normozoospèrmics amb infertilitat d'origen desconegut.

Metodologia: Es van utilitzar les mostres de 27 pacients que se sotmetien a FIV-ICSI, en funció de la qualitat embrionària obtinguda en dia 2 de desenvolupament embrionari, dividint embrions de bona qualitat (% A+B) i embrions de baixa qualitat (% D). Els pacients inclosos a l'estudi van ser classificats en tres subgrups segons els paràmetres seminals i les característiques fenotípiques. Les proteïnes van ser extretes, digerides i marcades mitjançant TMT10plex i subseqüentment es van analitzar per cromatografia líquida bidimensional seguit d'espectrometria de masses en tàndem (2D-LC-MS/MS).

Resultats: Es van identificar un total de 1409 proteïnes amb almenys un pèptid únic i després d'aplicar uns criteris de quantificació estrictes i fiables, on només es té en consideració per les anàlisis posteriors aquells pèptids quantificats per ≥2 *peptide-spectrum matches* (PSMs) en totes les mostres amb un coeficient de variació < 50 % en almenys el 75 % de les mostres, 210 van ser utilitzades per a la comparació quantitativa entre mostres. En el subgrup de parelles amb infertilitat d'origen desconegut, 2 proteïnes es trobaven correlacionades amb la bona qualitat embrionària (% A+B), mentre que 16 es trobaven correlacionades amb la baixa qualitat embrionària (% D). Notablement, un percentatge elevat de les proteïnes correlacionades amb la baixa qualitat embrionària estaven associades a l'alteració del *chaperonin TCP-1 ring complex* (TRiC), que presenta un paper important en el plegament de moltes proteïnes. Algunes d'aquestes correlacions van poder ser validades o bé a través de dades disponibles derivades d'estudis similars publicats prèviament o bé a través de noves determinacions fent servir les mostres amb factor masculí o femení conegut. A més, 7 / 18 proteïnes van ser quantificades amb més d'un pèptid i es va realitzar la seva respectiva correlació amb la qualitat embrionària a nivell de pèptid. Mentre que 2 / 7 proteïnes tenien tots els pèptids correlacionats amb la qualitat embrionària, en les 5 restants no tots els pèptids es trobaven correlacionats. Aquesta heterogeneïtat observada a nivell de pèptid es podria explicar per la presència de modificacions posttraduccionals (PTMs) o isoformes de les proteïnes, plantejant un nou repte en estudis futurs de proteòmica quantitativa.

Conclusions: Els resultats d'aquest estudi donen suport a que el proteoma de l'espermatozoide presenta un paper en l'embriogènesi i alguna de les proteïnes d'aquest estudi podrien acabar resultant útils com a potencials biomarcadors de predicció del resultat de tècniques de reproducció assistida.

RBMO

ELSEVIER

ARTICLE

Sperm proteomic changes associated with early embryo quality after ICSI

BIOGRAPHY

Meritxell Jodar participates in the EUGIN-UB Research Excellence Program and is a postdoctoral researcher at Fundació Clínic and Associate Professor at the University of Barcelona. Her main research focus is the molecular study of spermatozoa by different omic technologies to develop molecular clinical biomarkers and to improve male infertility diagnosis and prognosis.

Meritxell Jodar[1,2,*], Claudio Attardo-Parrinello[1,2,*], Ada Soler-Ventura[1,2,*], Ferran Barrachina[1,2], David Delgado-Dueñas[1,2], Salvadora Cívico[1,4], Josep Maria Calafell[1,4], Josep Lluís Ballescà[1,4], Rafael Oliva[1,2,3,]**

KEY MESSAGE

Sperm proteins seem to play a role in the correct development of the preimplantation embryo such as proteins of chaperonin-containing T-complex required for the correct folding of 10% of eukaryotic proteins. Those sperm proteins could be further developed as relevant clinical biomarkers to the better counselling of idiopathic infertile couples.

ABSTRACT

Research question: Do alterations of human sperm protein profile affect embryo quality?

Design: Sperm proteins from 27 infertile couples undergoing intracytoplasmic sperm injection (ICSI) were extracted and digested. The resulting peptides were labelled using tandem mass tags, separated by two-dimensional liquid chromatography, and identified and quantified using tandem mass spectrometry. Subsequently, sperm protein and peptide abundance were statistically analysed for correlation with ICSI-derived embryo quality in the subset of idiopathic infertile couples. Detected correlations were further assessed in the subset of infertile patients with a known factor.

Results: The abundance of 18 individual sperm proteins was found to correlate with embryo quality after ICSI. Of note, a high percentage of poor-quality ICSI-derived embryos was associated with alterations in several components of the eight-membered chaperonin-containing T-complex, which plays an important role in the folding of many essential proteins. Additionally, the abundance of sperm proteins with known functions in embryogenesis, such as RUBVL1, also correlated with early embryo quality ($r = -0.547$; $P = 0.028$). Some of the correlations found in this study were validated using either proteomic data from infertile patients with a known factor or data from similar published studies. Analysis at the peptide level revealed the association of some correlations with specific post-translational modifications or isoforms.

Conclusions: Our results support the hypothesis that the sperm proteome plays a role in early embryogenesis. Moreover, several sperm proteins have emerged as potential biomarkers that could predict the outcome of in-vitro assisted reproductive technologies, leading to the possibility of improved diagnosis of couples with idiopathic infertility.

[1] EUGIN-UB Research Excellence Program
[2] Molecular Biology of Reproduction and Development Research Group, Fundació Clínic per a la Recerca Biomèdica (FCRB), Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, University of Barcelona, Casanova 143, Barcelona 08036, Spain
[3] Biochemistry and Molecular Genetics Service, Hospital Clínic, Villarroel 170, Barcelona 08036, Spain
[4] Clinic Institute of Gynaecology, Obstetrics and Neonatology (ICGON), Hospital Clínic, Villarroel 170, Barcelona, Spain
*Contributed equally to the work.

KEYWORDS
Embryo quality
ICSI
Idiopathic infertility
Sperm proteome

**Corresponding author. E-mail address: roliva@ub.edu (R Oliva). https://doi.org/10.1016/j.rbmo.2020.01.004 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

INTRODUCTION

Infertility is an increasing health problem, affecting 20% of couples of reproductive age and directly affecting a patient's quality of life (*Skakkebaek et al., 2015*). Couples who experience infertility, defined as being unable to achieve pregnancy after 12 months of unprotected intercourse, are advised to seek fertility treatment (*Skakkebaek et al., 2015*). Although male and female factors contribute about equally to infertility (*Krausz et al., 2015*), most current clinical studies of infertile couples are focused on the woman. The evaluation of infertility in men is mainly restricted to the examination of semen parameters, which only reveals gross deficiencies in sperm count, motility and morphology, and therefore has a limited ability to predict the success of assisted reproductive technologies (ART) (*World Health Organization, 2010*).

Today, about 2.1% of children born in Europe are conceived through ART (*De Geyter et al., 2018*). Couples, however, often face several ART cycles, a costly and time-consuming process. Nowadays, intracytoplasmic sperm injection (ICSI), in which a single spermatozoon is injected directly into a mature oocyte, is widely used in-vitro fertilization (IVF) clinics. It is used in almost one-half of IVF patients and results in a pregnancy success rate of 27% per cycle (*De Geyter et al., 2018*). Although ICSI was initially recommended to specifically treat infertile couples with a known male factor, its use has become increasingly popular even for those couples with idiopathic infertility, in whom no male or female factor has been diagnosed. This has arisen because of increasing expectations of successful pregnancy from infertile couples. Some idiopathic infertile couples, however, fail to conceive after several ICSI cycles. *A priori*, unsuccessful in-vitro ART cycles are attributed to alterations in the oocyte as, traditionally, oocyte cytoplasm has been thought to contain all the factors required for early embryogenesis (*Swain and Pool, 2008*). Failed ICSI cycles, however, have also been observed in idiopathic infertile patients using donor oocytes, suggesting that some male molecular factors could contribute to the developing embryo (*Goudakou et al., 2012*). Indeed, several studies have revealed that spermatozoa provide DNA containing different epigenetic marks in combination with a complex population of proteins and RNAs that might be crucial for embryonic development (*Hammoud et al., 2009*; *Jodar et al., 2013*; *Carrell et al., 2015*; *Castillo et al., 2018*). These findings emphasize the need to develop alternative strategies for accurate evaluation of male factors potentially involved in the first steps of embryogenesis.

A total of 6871 non-redundant proteins have so far been identified in human spermatozoa using high-throughput proteomic strategies (*Martínez-Heredia et al., 2006*; *de Mateo et al., 2011*; *Amaral et al., 2013*; *2014a*; *Baker et al., 2013*; *Wang et al., 2013*; *Castillo et al., 2018*). Although most of these proteins are related to spermatogenesis and sperm functionality (*Amaral et al., 2014a*; *2014b*; *Castillo et al., 2018*), 11% of sperm proteins seem to be involved in fertilization, early embryo development or regulation of gene expression after embryonic genome activation (*Castillo et al., 2018*). Recently, integrative analyses of the human spermatozoon, oocyte, and embryo proteomes and transcriptomes have revealed a set of embryo proteins that may be exclusively paternal in origin. Some of these seem to be crucial for early embryogenesis, such as the centrosomal protein 135 (CEP135) (*Castillo et al., 2018*). Supporting this hypothesis, several studies have found strong correlations between alterations in the abundance of specific sperm proteins and low fertilization rates, blastocyst quality, as well as unsuccessful IVF or ICSI cycles (*Zhu et al., 2013*; *Azpiazu et al., 2014*; *Frapsauce et al., 2014*; *Légaré et al., 2014*; *McReynolds et al., 2014*; *Liu et al., 2018*). Except for the study published by *Zhu et al. (2013)*, these comparative proteomic studies were conducted using pools of samples categorized according to in-vitro ART outcomes. This strategy of pooling samples assumes that the underlying cause for the ART treatment outcome is the same in each individual sample. It is known, however, that several variables, including underlying unknown male, female factors, or both, could result in low fertilization rates, low blastocyst quality or unsuccessful ART cycles. Therefore, the use of individual samples could help to detect sample-specific alterations contributing to an unsuccessful ART cycle.

The aim of this study was to identify sperm proteins that could have a role during the first steps of embryogenesis. To achieve the objective, we correlated the abundance of sperm proteins with the quality of embryos derived from an ICSI cycle from infertile patients without any known male or female factor (idiopathic infertile patients) who underwent ICSI. Our results provide insights into the role of specific sperm proteins in early embryogenesis, which could be further developed as potential clinical biomarkers to predict the success of in-vitro ART.

MATERIALS AND METHODS

Sample collection and embryo quality assessment

Human semen samples were obtained from 27 infertile patients undergoing ICSI from the Assisted Reproduction Unit of the Hospital Clínic of Barcelona (FIV Clínic-ICGON, Spain). All human samples were used in accordance with the appropriate ethical guidelines and Internal Review Board guidelines of the Hospital Clínic of Barcelona (Spain; International Review Board study number R121031-096-2012/7942). International Review Board approval was granted on 22 November 2012. Written informed consent were obtained from all patients in accordance with the Declaration of Helsinki.

Semen samples were collected by masturbation into a sterile container after 2–7 days of sexual abstinence. Patients were classified according to their seminal parameters in line with World Health Organization guidelines (*World Health Organization, 2010*) and phenotypic traits as follows: idiopathic infertile couples (n = 16) (TABLE 1); infertile couples with a known male factor according to their seminal parameters (n = 4) (TABLE 1), including men with asthenozoospermia (<32% sperm progressive motility) or men with oligoasthenozoospermia (<39 million of total sperm and <32% sperm progressive motility); and infertile couples with a known female factor, including women aged over 40 years, with obesity according to body mass index, or both (n = 7) (TABLE 1).

Spermatozoa were purified from seminal plasma through a discontinuous density gradient (90%, 70%, 50% PureSperm) (Nidacon Laboratories AB, Gothenburg, Sweden) by centrifugation at 400 g for 20 min. Pelleted sperm were carefully recovered by direct aspiration of the

TABLE 1 SAMPLE CHARACTERISTICS SHOWING SEMEN PARAMETERS FROM DIAGNOSTIC SEMINOGRAM, PHENOTYPIC TRAITS AND REPRODUCTIVE CLINICAL INFORMATION

						ART data							
		Diagnostic seminogram				Embryo quality at day 2							
Sample number	Male age (years)	Sperm concentration (10^6 sperm/ml)	Sperm progressive motility (% a + b)[a]	Sperm morphology (% NF)[b]	Female age (years)	Recovered oocytes	MII Oocytes[c]	Fertilization rate (%)	A + B (%)[d]	C (%)[e]	D (%)[f]	Total embryos analysed	Pregnancy[g]
Idiopathic infertility													
1	42	>15	>32.0	>4	36	5	4	75.0	0.0	66.6	33.3	3	N
2	43	>15	>32.0	>4	<35[h]	16	16	68.8	9.1	63.6	27.3	11	N
3	39	>15	>32.0	>4	38	7	7	85.7	16.7	50.0	33.3	6	N
4	40	>15	>32.0	>4	33	6	6	100.0	16.7	83.3	0.0	6	N
5	38	>15	>32.0	>4	39	8	8	62.5	20.0	20.0	60.0	5	N
6	34	>15	>32.0	>4	34	7	7	42.9	33.3	66.6	0.0	3	N
7	37	>15	>32.0	>4	37	9	7	85.7	50.0	33.3	16.7	6	N
8	39	>15	>32.0	>4	37	8	7	100.0	57.1	28.5	14.3	7	Y
9	37	>15	>32.0	>4	36	8	6	100.0	66.6	16.6	16.7	6	Y
10	37	>15	>32.0	>4	32	10	8	75.0	66.6	16.6	16.7	6	N
11	37	>15	>32.0	>4	38	8	6	50.0	66.7	33.3	0.0	3	Y
12	32	>15	>32.0	>4	30	5	5	60.0	66.7	33.3	0.0	3	N
13	39	>15	>32.0	>4	<35[h]	9	7	100.0	71.4	14.3	14.3	7	Y
14	38	>15	>32.0	>4	39	8	8	75.0	83.3	16.6	0.0	6	Y
15	37	>15	>32.0	>4	37	5	4	75.0	100.0	0.0	0.0	3	Y
16	36	>15	>32.0	>4	36	8	7	100.0	100.0	0.0	0.0	7	Y
Male factor													
1	36	<15	<32.0	>4	34	9	6	83.3	20.0	80.0	0.0	5	Y
2	38	<15	<32.0	>4	36	16	13	61.5	42.0	14.0	42.0	7	N
3	39	>15	<32.0	>4	39	15	7	85.7	66.6	16.7	16.6	6	Y
4	39	>15	<32.0	>4	38	17	13	84.6	81.8	18.1	0.0	11	N
Female factor													
1	44	>15	>32.0	>4	41	7	5	60.0	33.3	33.3	33.3	3	N
2	36	>15	>32.0	>4	45	6	5	83.3	40.0	20.0	40.0	5	N
3	39	>15	>32.0	>4	41[i]	10	9	66.7	66.6	16.7	16.6	6	N
4	39	>15	>32.0	>4	39[i]	14	6	100.0	66.6	33.3	0.0	6	N
5	40	>15	>32.0	>4	40	5	5	80.0	75.0	25.0	0.0	4	Y
6	42	>15	<32.0	>4	42	8	7	85.7	100.0	0.0	0,0	6	N
7	40	>15	>32.0	>4	40	12	9	33.3	100.0	0.0	0,0	3	Y

[a] % (a + b): sperm progressive motility percentage.
[b] % Normal forms (NF) percentage.
[c] Metaphase II (MII) oocytes.
[d] % (A+B): top quality (A) and good but not suitable for elective single embryo quality transfer (B) percentage.
[e] % C: impaired embryo quality percentage.
[f] % D: embryos not recommended for transfer percentage.
[g] Pregnancy: presence (Y) or absence (N) of fetal heartbeats.
[h] Oocyte donation.
[i] Body mass index greater than 30.

pellet from the bottom of a conical tube, thus reducing possible contamination from overlaying layers. Afterwards, the pellet was resuspended in All grad Wash® media (The LifeGlobal® Group, Connecticut, USA) and centrifuged at 400 g for 10 min. Finally, sperm pellets were prepared by the routine swim-up procedure. Briefly, sperm pellets were gently covered with 0.5 ml of global® total® for Fertilization medium (The LifeGlobal® Group, Connecticut, USA) and incubated for 20 min at 37°C in 7% of CO_2 and 7% of O_2, allowing motile spermatozoa to swim into the medium. A small aliquot of spermatozoa that migrated to the top of the medium was transferred to a new

dish, and qualified operators selected the best spermatozoa according to sperm morphology. These sperm cells were used for the ICSI procedure. The surplus of sperm cells not used for ICSI was assessed for absence of round cells by optical microscopy, incubated with CryoSperm (Origio, Malow, Denmark) following manufacturer's instructions, and cryopreserved in liquid nitrogen until further processing for proteomic analysis (FIGURE 1).

Embryo quality assessment (n = 150) was conducted after the ICSI procedure in accordance with the standard assisted reproduction clinic protocol. Briefly, fertilized oocytes were examined and only those with two pronuclei were further cultured in 20 µl drops of global® culture media (The LifeGlobal® Group, Guilford, CT, USA) under mineral oil. Embryos at day 2 were scored based on the morphological features according to the embryo assessment criteria established by the Spanish Society, Asociación Española para el Estudio de la Biología de la Reproducción (ASEBIR) (*Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group Of Embryology, 2011*). The embryo quality assessment was mainly based on six parameters: the number of blastomeres; symmetry of the blastomeres; size and distribution of the embryo fragmentation; presence of multi-nucleation; cytoplasmic anomalies such as the presence of vacuoles; and zona pellucida irregularities. The ASEBIR consensus scheme scores day-2 embryos as (A) top quality, (B) good-quality but not suitable for elective single-embryo transfer, (C) impaired embryo quality and (D) not recommended for transfer. Our patients were classified according to the percentage of good-quality embryos (embryos graded as A or B) or poor-quality embryos (embryos graded as D), which are not recommended for transfer after an ICSI cycle (TABLE 1). Finally, as established in the standard protocol of the clinic, one or two good-quality embryos were transferred to women on day 3. The clinical outcome 'pregnancy' was evaluated by the presence of a fetal heartbeat.

Protein purification and peptide tandem mass tag-labelling

Sperm proteins were solubilized by incubation of spermatozoa for 1 h at room temperature with 1% sodium dodecyl sulphate in 100 mM triethylammonium bicarbonate. Cell lysates were centrifuged at 17,500 g for 10 min at 4°C and the supernatants containing the soluble protein fraction were quantified using Pierce™ Reducing Agent Compatible BCA Protein Assay Kit (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions. From each sample, 30 µg of proteins were used for tandem mass tag (TMT)-10 plex labelling by following the manufacturer's instructions with minor modifications, as described elsewhere (*Bogle et al., 2017*). Before peptide labelling, equivalent aliquots of each individual sample (n = 27) were merged in a single internal control. Subsequently, each sample and the internal control were individually labelled with TMT reporter ions with intensities from mass-to-charge ratios (m/z) 126:131; TMT-126 was used for the internal control and –127N, –127C, –128N, –128C, –129N, –129C, –130N, –130C, –131 for individual samples. To quench the reaction, each tube was incubated with 2.6 µl of 8% hydroxylamine for 15 min. After quenching, TMT-labelled samples were combined at equal volume in three different pools (A, B and C) with nine samples and the internal control each. The pooled samples were dried in a speed-vacuum and, subsequently, 20 µg of peptides from each pool were cleaned-up with Pierce C18 Spin Columns (Thermo Scientific, Waltham, MA, USA), according to the manufacturer's instructions.

Two-dimensional liquid chromatography with tandem mass spectrometry

A total of 10 µg of TMT-labelled peptides from each pool (A, B and C) were fractioned using a Zorbax 300 HILIC column (1.8 µm, 2.1 × 50 mm, 300 Å) (Agilent Technologies, Santa Clara, CA, USA) coupled with a Surveyor MS Pump (Thermo Finnigan, Waltham, MA, USA). Two different buffers were used for fractionation: Buffer A (97% acetonitrile [ACN], 3% H2O) and Buffer B (97% 50 mM ammonium acetate pH = 4.5, 3% ACN). An elution gradient from 5% to 100% buffer B was run over 15 min, and 20 fractions were collected from each pool. Each fraction was dried under a speed vacuum, reconstituted and loaded to a PepMap 100 trap column (300 µm x 5 mm, 5 µm, 100 Å) (Thermo Fisher Scientific, Waltham, MA, USA) to clean and pre-concentrate the fractions. For the liquid chromatography, a PepMap was used with a 120 min linear gradient with buffer A (97% H_2O, 3% ACN, 0.1% formic acid [FA]) and buffer B (97% ACN, 3% H_2O, 0.1% FA) as follows: 0–4 min, 0% buffer B; 4–100 min, 32.5% buffer B; and 100–120 min, 100% buffer B. The tandem mass spectrometry analysis was carried out using a nano-LC ultra 2D eksigent (AB Sciex, Madrid, Spain) attached to an LTQ-Orbitrap Velos coupled with a nanospray ion source (Thermo Fisher Scientific, Waltham, MA, USA). The LTQ Orbitrap Velos (Thermo Fisher Scientific, Waltham, MA, USA) settings included 30,000 of resolving power (full width half medium) for mass spectrometry level 1 scans at 400 m/z for precursor ions followed by 30,000 of resolving power (full width half medium) for mass spectrometry level 2 scans of the 20 most intense precursor ions at 400 m/z, in positive ion mode. The mass range was set 300–1600 m/z. The acquisition of tandem mass spectrometry data was conducted using Xcalibur 2.2 (Thermo Fisher Scientific, Waltham, MA, USA). Normalized collision energy for higher energy collisional dissociation (HCD)-MS2 was set to 42%.

Protein identification and quantification

Tandem mass spectrometry data output was analysed by Proteome Discoverer 1.4. (ThermoFisher Scientific, Waltham, MA, USA). For database searching, processed data were submitted to the in-house Homo sapiens UniProtKB/Swiss-Prot database (last modified February 2014; 20,240 protein entries) using SEQUEST version 28.0 (Thermo Fisher Scientific, Waltham, MA, USA). Percolator search node was also applied. Searches were conducted using the following settings: two maximum missed cleavages for trypsin; TMT-labelled in N-terminus, lysine, histidine, serine, tyrosine (+229.163 Da), and methionine oxidation (+15.995 Da) as dynamic modifications, and cysteine carbamidomethylation (+57.021 Da) as a static modification; 20 ppm precursor mass tolerance, 0.05 Da fragment mass tolerance, 5 mmu peak integration tolerance, and most confident centroid peak integration method.

Criteria used to accept protein identification included a minimum of one unique peptide matched per protein, with a 1% false discovery rate. The 'enable protein grouping' option was disabled to avoid possible ambiguity in the identification of different isoforms of the same proteins during

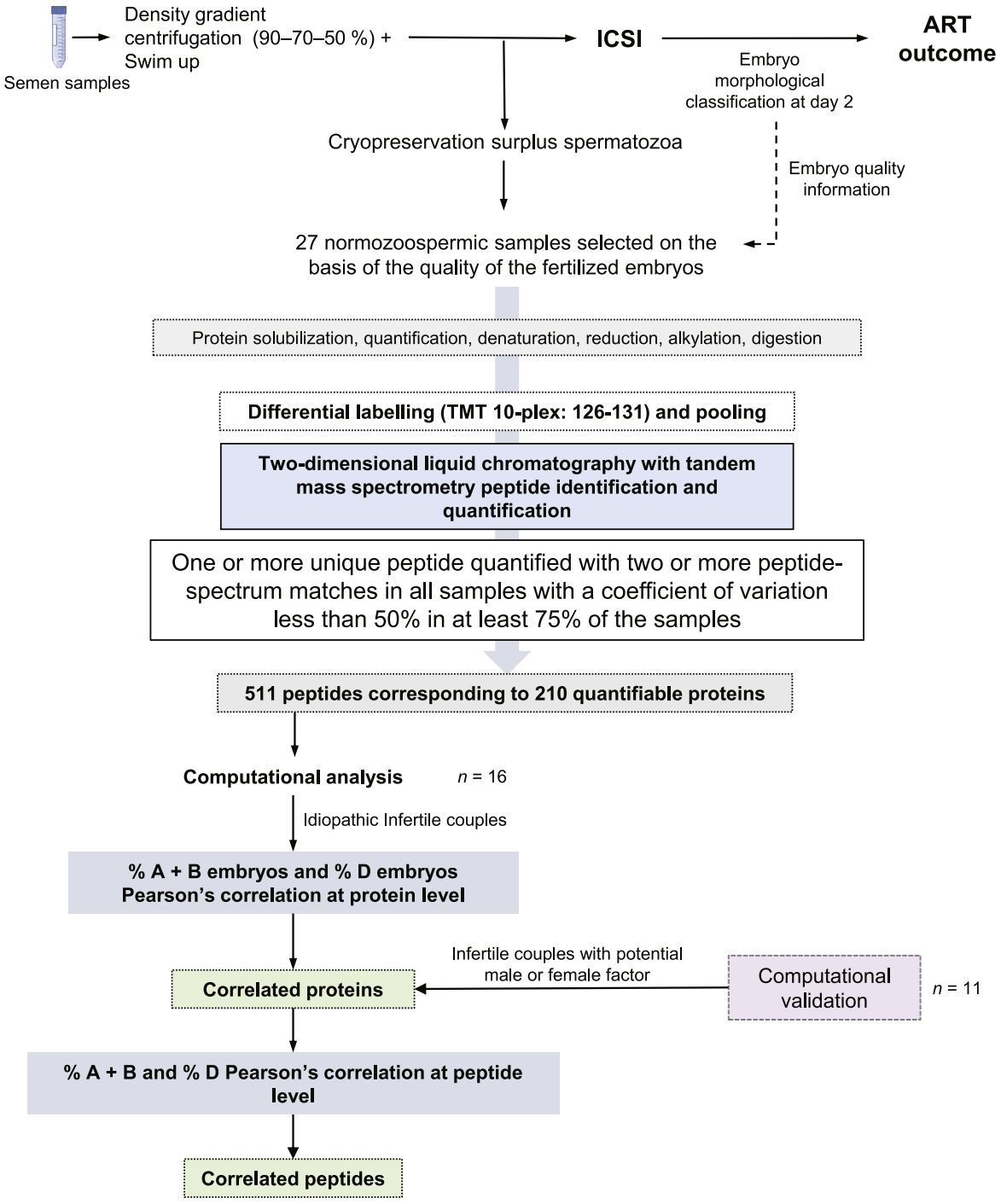


FIGURE 1 Study design. Overall procedure used for the identification and quantification of proteins in the different types of patients through tandem mass tag-10plex peptide labelling followed by two-dimensional liquid chromatography with tandem mass spectrometry, and subsequent computational analysis. ART, assisted reproductive technology; ICSI, intracytoplasmic sperm injection; TMT, tandem mass tag.

quantification. For protein quantification data, normalized TMT quantitative values expressed as the ratio of the intensities of reporter ions from HCD MS2 spectra from each individual sample (TMT-127 to TMT-131) with the internal control (TMT-126) were obtained by Proteome Discoverer software from each identified spectrum (Supplementary Figure 1), as previously described (Barrachina et al., 2019). Only those peptides quantified by ≥ 2 peptide-spectrum matches (PSMs) in all the samples with a coefficient of variation < 50% in at least 75% of the samples were considered for further statistical analyses (*Barrachina et al., 2019*). Significant statistical differences between groups were evaluated after the normalization of the relative proteomic quantification values by $\log_2$ transformation.

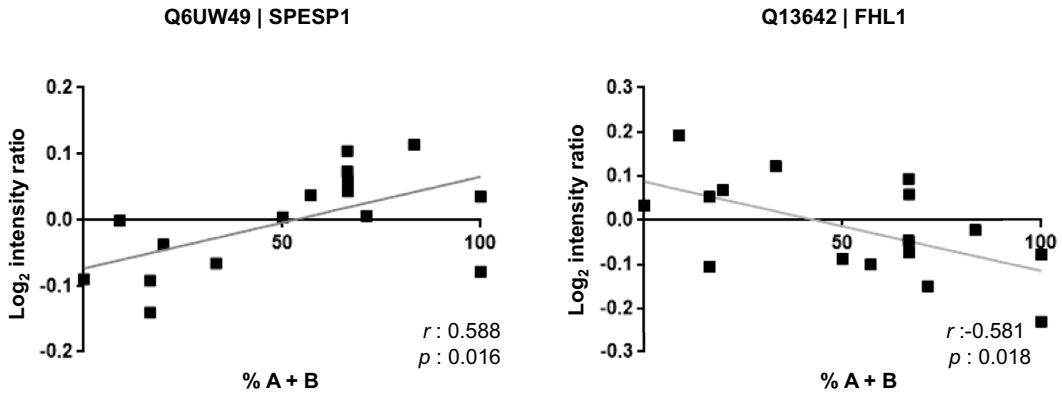


FIGURE 2 Correlation between sperm protein abundance and percentage of good-quality embryos in idiopathic infertile patients. (A) Scatter plot of normalized SPESP1 protein abundance generated by two-dimensional liquid chromatography with tandem mass spectrometry (log2 intensity ratio of the proteins) versus per cent good-quality embryos (A + B) derived from intracytoplasmic sperm injection; (B) scatter plot of normalized FHL1 protein abundance generated by two-dimensional liquid chromatography with tandem mass spectrometry (log2 intensity ratio of the proteins) versus per cent good-quality embryos (A + B) derived from intracytoplasmic sperm injection. The Pearson correlation coefficient (r) and its *P*-value are indicated for both correlations.

Statistical analyses

R software version 3.4.4 (http://www.r-project.org) or GraphPad Prism software version 7.01 (GraphPad Software Inc., San Diego, USA) were used for all data analyses. Multiple Pearson correlation analyses between $\log_2$ transformed values of relative proteomic quantification at both protein and peptide levels with either % A + B or % D embryos were separately conducted for the three groups: idiopathic infertile couples; infertile couples with a male factor; and infertile couples with a female factor. $P < 0.05$ was considered statistically significant. Additionally, the UniProt Knowledgebase (https://www.uniprot.org/) was used to identify possible protein isoforms and post-translational modifications.

To infer the functional involvement of the sperm proteins correlated with embryo quality, gene ontology enrichment analysis was carried out using Gene Ontology Consortium database (http://www.geneontology.org/) based on PANTHER v. 14 database (Released on 2019-01-01). This enrichment analysis finds which gene ontology terms are over-represented using functional annotations from the Gene Ontology Consortium database for target sperm proteins. The significance of the enrichment was calculated by Fisher's exact test. $P < 0.05$ after Bonferroni adjustment was considered statistically significant.

RESULTS

Identification and reliable quantification of sperm proteins

In the present study, 1409 proteins were identified, of which 727 were detected in all the 27 sperm samples assessed (Supplementary Table 1). From those, 511 peptides derived from 210 proteins met our strict quantification criteria (peptides with ≥2 PSMs quantified and with a coefficient of variation <50% in at least 75% of the samples), and were used for further correlation analyses (FIGURE 1 and Supplementary Table 2). Gene ontology enrichment analysis showed that this subset of proteins is involved in processes such as energy production, spermatogenesis, sperm motility, fertilization and telomere maintenance (Supplementary Table 3).

Correlation of sperm proteins from idiopathic infertile couples with embryo quality after ICSI

The relative abundance of two sperm proteins, the sperm equatorial segment protein 1 (SPESP1) and the four and a half LIM domains protein 1 (FHL1), displayed significantly positive ($r = 0.588$, $P = 0.016$) and negative correlations ($r = -0.581$; $P = 0.018$), respectively with the percentage of ICSI-derived good-quality embryos (classed as A+B) in idiopathic infertile patients (FIGURE 2). Additionally, the abundance of 16 sperm proteins was significantly correlated with the percentage of ICSI-derived poor-quality embryos (classed as D) (TABLE 2; *P*-values shown in TABLE 2, all <0.05). Whereas 15 of the 16 sperm proteins showed a negative correlation with the percentage of poor-quality embryos, just one sperm protein, the uncharacterized protein C7orf61, displayed a positive correlation.

Correlations between sperm proteins at the peptide level from idiopathic infertile couples and embryo quality

Up to seven of the 18 sperm proteins whose relative abundance correlated with embryo quality were reliably quantified with more than one peptide. Specifically, these proteins are the protein T-complex protein 1 subunit gamma (CCT3), the protein T-complex protein 1 subunit theta (CCT8), the enolase 1 (ENO1), the zona pellucida binding protein (ZPBP), the mitochondrial protein peroxiredoxin-5 (PRDX5), RuvB-like 1 (RUVBL1) and SPESP1 (TABLE 3). Pearson correlation analysis at the peptide level showed that the abundance of one or more peptides of all proteins assessed did not correlate with embryo quality (TABLE 3). This heterogeneity observed at the peptide level may be explained by the presence of either post-translational modifications (PTM) or protein isoforms (FIGURE 3). For instance, only two out of three peptides quantified for CCT3 showed a strong correlation with poor-quality embryos. Interestingly, the non-correlated peptide contained a threonine (T) residue that could be a target of phosphorylation (FIGURE 3A). Similarly, only one out of

TABLE 2 PEARSON CORRELATION ANALYSIS BETWEEN NORMALIZED RELATIVE SPERM PROTEIN ABUNDANCE AND PERCENTAGE OF POOR-QUALITY EMBRYOS (D)[a]

Gene name	*UniprotKB*	*Pearson correlation coefficient (r-value)*	*Significance (P-value)*
C7orf61	Q8IZ16	+0.502	0.048
CCT7	Q99832	–0.646	0.007
CCT2	P78371	–0.641	0.007
CCT3	P49368	–0.626	0.010
CCT8	P50990	–0.622	0.010
ENO1	P06733	–0.587	0.017
CA2	P00918	–0.580	0.019
PRDX6	P30041	–0.560	0.024
RUVBL1	Q9Y265	–0.547	0.028
TCP1	P17987	–0.544	0.030
CRISP2	P16562	–0.541	0.030
ZPBP	Q9BS86	–0.520	0.039
GLB1L	Q6UWU2	–0.516	0.041
DPCD	Q9BVM2	–0.510	0.043
PSMA6	P60900	–0.507	0.045
PRDX5	P30044	–0.499	0.049

Normalized relative sperm protein abundance assessed by two-dimensional liquid chromatography with tandem mass spectrometry data.

[a] Embryo assessment criteria established by the Spanish Society. Asociación Española para el Estudio de la Biología de la Reproducción (ASEBIR) (*Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group Of Embryology. 2011*).

two peptides quantified for the PRDX5 was correlated with the percentage of D embryos. The presence of different isoforms, however, produced by alternative splicing could explain this observation. In this case, whereas the non-correlated peptide is common in all PRDX5 isoforms provided by UniProt Knowledgebase, the correlated peptide is exclusively absent in PRDX5 isoform 4. Therefore, these results suggest that the abundance of all the PRDX5 isoforms present in sperm except for the isoform 4 are correlated with embryo quality (FIGURE 3B).

Validation of the identified correlations between sperm proteins and embryo quality in a different subset of infertile patients

The 18 sperm proteins correlated with embryo quality in idiopathic infertile couples were assessed in a different subset of infertile patients with a known male or female factor (n = 11) for its validation. Although a significant positive correlation was observed between the abundance of the uncharacterized protein C7orf61 and the percentage of D embryos in infertile patients with known male factor (r = 0.99; P = 0.01), no correlations were found in patients with a known female factor (Supplementary Table 4), as was expected.

DISCUSSION

Traditionally, women have been the focus of reproductive medicine. The investigation of the male partner has been largely restricted to evaluation of seminal parameters. This is reflected in the fact that, currently, no pharmaceutical compounds are available to prescribe to infertile males, or to improve sperm quality *in vitro*. Similarly, validated molecular tests based on the study of semen to predict the success of ARTs are lacking. The lack of tools currently available for male infertility diagnosis, prognosis and therapy is mainly due to the incomplete understanding of the father's contribution to his progeny. Deciphering what molecular causes are involved in male infertility may shed new light on paternal components essential for embryogenesis and may promote the development of molecular tests that would finally help to change the male fertility diagnosis paradigm (*Jodar et al., 2015*).

The use of comparative high-throughput proteomics of individual sperm samples is a good strategy for the discovery of new physiological and pathological mechanisms of male fertility, as well as clinically useful biomarkers (*Jodar et al., 2017*). To avoid false-positive results, however, it is essential to establish strict criteria for both protein identification and quantification. In the present study, 18 sperm proteins were correlated with ICSI-derived embryo quality. It is interesting that alterations in the abundance of seven of those 18 sperm proteins has been previously associated with in-vitro ART outcomes in similar published studies (Supplementary Table 5). This is exemplified by the strong association between blastocyst development and the differential abundance of cysteine-rich secretory protein 2 (CRISP2) and PRDX5 (*McReynolds et al., 2014*). Similarly, the abundance of FHL1, RUVBL1, ENO1, C7orf61, and ZPBP in spermatozoa was related to the pregnancy outcome through conventional IVF or ICSI (*Azpiazu et al., 2014*; *Légaré et al., 2014*; *Liu et al., 2018*).

TABLE 3 PEARSON CORRELATION ANALYSIS BETWEEN EMBRYO QUALITY (% A + B EMBRYOS OR % D EMBRYOS) AND NORMALIZED RELATIVE SPERM PEPTIDE ABUNDANCE (TWO-DIMENSIONAL LIQUID CHROMATOGRAPHY WITH TANDEM MASS SPECTROMETRY) FOR THE SEVEN SPERM PROTEINS IDENTIFIED AS CORRELATED AT PROTEIN LEVEL AND QUANTIFIED WITH MORE THAN ONE PEPTIDE. PEARSON CORRELATION COEFFICIENT (R-VALUE) AND *P*-VALUE ARE PROVIDED.

% A+B

UniprotKB	*Gene name*	*Peptide sequence and modifications*	*Pearson's correlation (r-value)*	*P-value*
Q6UW49	*SPESP1*	aATVFNTLk N-Term (TMT10-plex); K9 (TMT10-plex)	+0.5661	**0.0222**
		sPVTTLDk S1 (TMT10-plex); K8 (TMT10-plex)	+0.4769	*0.0618*
		sQLLPVGR S 1(TMT10-plex)	+0.5014	**0.0478**
		Protein level	+0.588	0.0166

% D

Protein	*Gene name*	*Peptide sequence and modifications*	*Pearson's correlation (r-value)*	*P-value*
P49368	*CCT3*	iPGGIIEDScVLR N-Term (TMT10-plex); C10(Carbamidomethyl)	−0.5017	**0.0477**
		tLIQNcGASTIR N-Term (TMT10-plex); C6 (Carbamidomethyl)	−0.3572	0.1745
		vQSGNINAAk N-Term (TMT10-plex); K10 (TMT10-plex)	−0.6326	**0.0085**
		Protein level	−0.6256	**0.0095**
P50990	*CCT8*	aLAENSGVk N-Term (TMT10-plex); K9 (TMT10Plex)	−0.6596	**0.0054**
		aVDDGVNTFk N-Term (TMT10-plex); K10 (TMT10-plex)	−0.6368	**0.008**
		lATNAAVTVLR N-Term (TMT10-plex)	−0.3358	0.2035
		Protein level	−0.6225	**0.01**
P06733	*ENO1*	eGLELLk N-Term (TMT10-plex); K7 (TMT10-plex)	−0.5593	**0.0243**
		gNPTVEVDLFTSk N-Term (TMT10-plex); K13 (TMT10-plex)	−0.6453	**0.0069**
		lMIEMDGTENk N-Term (TMT10-plex); K11 (TMT10-plex)	−0.4831	*0.058*
		lNVTEQEk N-Term (TMT10-plex); K8 (TMT10-plex)	−0.4129	0.1119
		tIAPALVSk N-Term (TMT10-plex); K9 (TMT10-plex)	−0.1103	0.6842
		yISPDQLADLYk N-Term (TMT10-plex); K12 (TMT10-plex)	−0.5463	**0.0286**
		Protein level	−0.5866	0.0169
Q9Y265	*RUVBL1*	aVLLAGPPGTGk N-Term (TMT10-plex); K12 (TMT10-plex)	−0.4906	*0.0537*
		lDPSIFESLQk N-Term (TMT10-plex); K11 (TMT10-plex)	−0.5013	**0.0479**
		Protein level	−0.5474	0.0282
Q9BS86	*ZPBP*	aYVMLHQk N-Term (TMT10-plex); K8 (TMT10-plex)	−0.4429	*0.0858*
		fFNQQVEILGR N-Term (TMT10-plex)	−0.1008	0.7103
		iVGSTSFPVk N-Term (TMT10-plex); K10 (TMT10-plex)	−0.52	**0.039**
		nAELIDPSFQWYGPk N-Term (TMT10-plex); K15(TMT10-plex)	−0.5808	**0.0183**
		Protein level	−0.5199	0.039
P30044	*PRDX5*	lLADPTGAFGk N-Term (TMT10-plex); K11 (TMT10-plex)	0.04154	0.8786
		vNLAELFk N-Term (TMT10-plex); K8 (TMT10-plex)	−0.6012	**0.0138**
		Protein level	−0.4993	0.0489

TMT, tandem mass tag.

This concordance between our results and those studies published by others (around 40%) is much higher than that observed when comparing other phenotypes. For instance, only 9% of concordance has been observed between comparative high-throughput sperm proteomics studies of patients with asthenozoospermia (*Jodar et al., 2017*). This observation leads us to conclude that applying our strict criteria for protein identification and quantification is a good approach to reduce false-positive associations.

The abundance of only two sperm proteins was significantly correlated with the percentage of good-quality (A + B) ICSI-derived embryos in idiopathic infertile couples (FIGURE 2), a negative correlation for FHL1 and positive correlation with SPESP1, with the latter being essential for sperm-egg binding and fusion (*Fujihara et al., 2010*). Although the ICSI procedure overcomes the sperm-oocyte binding and fusion processes, it has been shown that *Spesp1* disruption in mice causes aberrant distribution of various sperm proteins, resulting in fertilized oocytes that do

A) Part of the peptides quantified in a protein does not correlate due to the presence of post-translational modifications

P49368 | *CCT3* (Correlated at the peptide level with % D embryos).

MMGHRPVLVLSQNTKRESGRK**VQSGNINAAK**TIADIIRTCLGPKSMMKMLLDPM…ACNIALDAVKMVQFEENGRKEIDIKKYAR
VEK**IPGGIIEDSCVLR**GVMINKDVTHPRMRRYIKNPRIVLLDSSLEYKKGESQTDIEITREEDFTRILQMEEEYIQQLCEDIIQ
LKPDVVI…ERNLQDAMQVCRNVLLDPQLVPGGGASEMAVAHALTEKSKAMTGVEQWPYRAVAQALEVIPR**TLIQNCGAS<u>T</u>IR**LL
TSLRAKHTQENCETWGV…LLLRIDDIVSGHKKKGDDQSRQGGAPDAGQE

P50990 | *CCT8* (Correlated at the peptide level with % D embryos).

MALHVPKAPGFAQMLKEGAKHF…LEEMGHCDSVYLSEVGDTQVVVFKHEKEDGAISTIVLRGSTDNLMDDIER**AVDDGVNTF<u>K</u>**V
LTRDKRLVPGGGATEIELAKQITSYGETCPGLEQYAIKKFAEAFEAIPR**ALAENSGV<u>K</u>**ANEVISKLYAVHQEGNKNVGLDIEAE
VPAVKDMLEAGILDTYLGKYWAIK**LATNAAVTVLR**VDQIIMAKPAGGPKPPSGKKDWDDDQND

B) Part of the peptides quantified in a protein does not correlate due to the quantification of an isoform or the canonical form of a protein.

P30044 | *PRDX5* (Correlated at the peptide level with % D embryos).

MGLAGVCALRRSAGYILVGGAGGQSAAAAAARRYSEGEWASGGVRSFSRAAAAMAPI<u>KVGDAIPAVEVFEGEPGNK</u>**VNLAELFK**<u>G
KKGVLFGVPGAFTPGCSKTHLPGFVEQAEALKAKGVQVVACLSVNDAFVTGEWGRAHKAEG</u>KVR**LLADPTGAFGK**ETDLLLDDS
LVSIFGNRRLKRFSMVVQDGIVKALNVEPDGTGLTCSLAPNIISQL

FIGURE 3 Correlations of sperm proteins with embryo quality at the peptide level. (A) Two of the three quantified peptides of CCT3 were correlated with the percentage of D embryos. Interestingly, the non-correlated peptide could contain a post-translational modification (a phosphorylated residue in T459) as described in the Uniprot Knowledgebase. In contrast, only the peptides from CCT8 that could contain PSMs (acetylated residue at K400 and a glycyl lysine isopeptide at K459) are correlated with percentage of D embryos. There is not any PTM described by Uniprot Knowledgebase for the non-correlated peptide of *CCT8* (B). Only one peptide of PRDX5 was correlated with the percentage of D embryos, most probably due to PRDX5-isoforms sequence differences. All isoforms of PRDX5 contain the non-correlated peptide; in contrast, the correlated peptide is missing in a specific isoform of PRDX5 (Isoform 4). Peptides with significant correlations are indicated in green ($P < 0.05$); (*Table 3*) for more details) and peptides that do not reach significance are indicated in red. Underlined and enlarged amino acid residues indicate the absence of a part of the amino acid sequence caused by an isoform or the presence of post-translationally modified residues.

not yield high-quality embryos (*Fujihara et al., 2010*).

Interestingly, 15 sperm proteins were negatively correlated and one positively correlated with the percentage of poor-quality ICSI-derived embryos (TABLE 2). According to the Gene Ontology Consortium Database, those 16 proteins are functionally enriched in processes related to binding of sperm to zona pellucida, protein folding and telomere maintenance among others (Supplementary Table 6). Of note, five of those 15 negatively correlated (t-complex protein 1 subunit; alpha (TCP1); eta (CCT7); beta (CCT2); gamma (CCT3) and theta (CCT8)) are components of the eight-membered chaperonin-containing T-complex (TRiC). This is essential for cell survival. It is responsible for the correct folding of 10% of the eukaryotic proteome (*Leitner et al., 2012*). Among the proteins that require TRiC for correct folding is Cajal body protein 1 (TCAB1), a telomerase cofactor which is essential for telomerase function (*Freund et al., 2014*). Therefore, the negative correlation among TRiC components abundance and the percentage of poor quality supports previous findings, which positively correlated the sperm telomere length with the quality of early embryonic development and the embryo morphology (*Turner et al., 2010*). No correlation was found, however, with pregnancy rates (*Yang et al., 2015*; *Torra-Massana et al., 2018*). Other sperm proteins, the abundance of which was also negatively correlated with the percentage of poor-quality embryos, have been described previously as essential in early embryogenesis. These include RUBVL1, PRDX5, and peroxiredoxin-6 (PRDX6). RUBVL1 is a chromatin remodeller potentially involved in the proliferation of the inner cell mass and, for this reason, *Rubvl1*-null embryos do not reach the blastocyst stage (*Bereshchenko et al., 2012*). In contrast, PRDX5 and PRDX6 have essential antioxidant activity during early embryogenesis (*Leyens et al., 2004*). The possible role of the remaining eight proteins that were correlated with poor-quality embryos in the first steps of embryogenesis has not yet been elucidated. Of note, however, four of them (CRISP2, C7orf61, ENO1, and ZPBP; Supplementary Table 5) have also been detected with an altered abundance in in-vitro ART outcomes, implicating their involvement in crucial events in the developing embryo (*Azpiazu et al., 2014*; *Légaré et al., 2014*; *Liu et al., 2018*).

The strong correlation between the abundance of C7orf61 protein with embryo quality is maintained in infertile patients with abnormal seminal parameters but lost in infertile couples with a known female factor, as expected. These results validate the potential predictive value of the sperm protein C7orf61 to estimate the quality of the in-vitro ART-derived embryos in infertile couples without any known female factor and regardless the presence of alterations in the seminal parameters. The validation of the predictive value of these sperm protein biomarkers in a larger number of infertile patients by targeted proteomics or other techniques, e.g. western blot or enzyme-linked immunosorbent assay, is now needed to establish their relevance and applicability to clinical practice. From results presented here, derived from proteomic analysis at the peptide level, post-translational modification (PTM) and different isoforms of the target proteins should be also taken into consideration before the design of further validation

tests. Additionally, the analysis of single sperm proteomic profile selecting sperm using the same criteria as are used for ICSI could also help to unravel the role of sperm proteome in early embryogenesis.

In conclusion, this high-throughput proteomic analysis of idiopathic infertile patients classified according to the quality of ICSI-derived embryos has thus provided insights into the contribution of the sperm to early embryogenesis and has suggested novel potentially useful biomarker leads for future clinical application. It is important to note that the negative correlation of several TRiC complex components with poor embryo quality supports the hypothesis that sperm telomeres are important for correct embryogenesis. Moreover, several sperm proteins have emerged as potential candidates to predict the success of in-vitro ART technologies, especially C7orf61 protein, as it seems to be associated with embryo quality regardless of any alterations in the seminogram. Taken together, these factors indicate that our results might be applicable to clinical practice, potentially leading to improvements in the diagnosis of infertile couples and the prediction of success in ART.

ACKNOWLEDGMENTS

This work was supported with funding from EUGIN-UB Research Excellence Program (EUREP 2014), from the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competividad; fondos FEDER 'una manera de hacer Europa' PI13/00699, PI16/00346), and from EU-FP7-PEOPLE-2011-ITN289880. M.J. is granted by Government of Catalonia (Generalitat de Catalunya, pla estratègic de recerca i innovació en salut, PERIS 2016-2020, SLT002/16/00337). F.B. is granted by Spanish Ministry of Education, Culture and Sports (Ministerio de Educación, Cultura y Deporte para la Formación de Profesorado Universitario, FPU15/02306). We greatly thank Dr Judit Castillo for critical revision of the manuscript. The authors also recognize Raquel Ferreti and Alicia Diez for their assistance in the routine seminograms and sample collection, and Dr Josep Maria Estanyol for his assistance in the proteomic workflow and analysis.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rbmo.2020.01.004.

REFERENCES

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group Of Embryology. **The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting.** Hum. Reprod. 2011; 26: 1270–1283

Amaral, A., Castillo, J., Estanyol, J.M., Ballescà, J.L., Ramalho-Santos, J., Oliva, R. **Human Sperm Tail Proteome Suggests New Endogenous Metabolic Pathways. Mol. Cell.** Proteomics 2013; 12: 330–342

Amaral, A., Castillo, J., Ramalho-Santos, J., Oliva, R. **The combined human sperm proteome: Cellular pathways and implications for basic and clinical science.** Hum. Reprod. Update 2014; 20: 40–62

Amaral, A., Paiva, C., Attardo Parrinello, C., Estanyol, J.M., Ballescà, J.L., Ramalho-Santos, J.J., Oliva, R. **Identification of Proteins Involved in Human Sperm Motility Using High-Throughput Differential Proteomics.** J. Proteome Res. 2014; 13: 5670–5684

Azpiazu, R., Amaral, A., Castillo, J., Estanyol, J.M., Guimerà, M., Ballescà, J.L., Balasch, J., Oliva, R. **High-throughput sperm differential proteomics suggests that epigenetic alterations contribute to failed assisted reproduction.** Hum. Reprod. 2014; 29: 1225–1237

Baker, M.A., Naumovski, N., Hetherington, L., Weinberg, A., Velkov, T., Aitken, R.J. **Head and flagella subcompartmental proteomic analysis of human spermatozoa.** Proteomics 2013; 13: 61–74

Barrachina, F., Jodar, M., Delgado-Dueñas, D., Soler-Ventura, A., Estanyol, J.M., Mallofré, C., Ballescà, J.L., Oliva, R. **Stable-protein Pair Analysis as A Novel Strategy to Identify Proteomic Signatures: Application To Seminal Plasma From Infertile Patients.** Mol. Cell. Proteomics 2019; 18: S77–S90

Bereshchenko, O., Mancini, E., Luciani, L., Gambardella, A., Riccardi, C., Nerlov, C. **Pontin is essential for murine hematopoietic stem cell survival.** Haematologica 2012; 97: 1291–1294

Bogle, O.A., Kumar, K., Attardo-Parrinello, C., Lewis, S.E.M., Estanyol, J.M., Ballescà, J.L., Oliva, R. **Identification of protein changes in human spermatozoa throughout the cryopreservation process.** Andrology 2017; 5: 10–22

Carrell, D.T., Aston, K.I., Oliva, R., Emery, B.R., De Jonge, C.J. **The "omics" of human male infertility: integrating big data in a systems biology approach.** Cell Tissue Res 2015; 363: 295–312

Castillo, J., Jodar, M., Oliva, R. **The contribution of human sperm proteins to the development and epigenome of the preimplantation embryo.** Hum. Reprod. Update 2018; 24: 535–555

De Geyter, C., Calhaz-Jorge, C., Kupka, M.S., Wyns, C., Mocanu, E., Motrenko, T., Scaravelli, G., Smeenk, J., Vidakovic, S., Goossens, V., Gliozheni, O., Strohmer, H., Petrovskaya, E., Tishkevich, O., Bogaerts, K., Balic, D., Sibincic, S., Antonova, I., Vrcic, H., Ljiljak, D., Pelekanos, M., Rezabek, K., Markova, M.J., Lemmen, J., Sõritsa, D., Gissler, M., Tiitinen, A., Royere, D., Tandler–schneider, A., Kimmel, M., Antsaklis, A.J., Loutradis, D., Urbancsek, J., Kosztolanyi, G., Bjorgvinsson, H., de Luca, R., Lokshin, V., Ravil, V., Magomedova, V., Gudleviciene, Z.,

Belo Lopes, G., Petanovski, Z., Calleja-Agius, J., Xuereb, J., Moshin, V., Simic, T.M., Vukicevic, D., Romundstad, L.B., Janicka, A., Laranjeira, A.R., Rugescu, I., Doroftei, B., Korsak, V., Radunovic, N., Tabs, N., Virant-Klun, I., Saiz, I.C., Mondéjar, F.P., Bergh, C., Weder, M., Smeenk, J.M.J., Gryshchenko, M., Baranowski, R. **ART in Europe, 2014: Results generated from European registries by ESHRE.** Hum. Reprod. 2018; 33: 1586–1601

de Mateo, S., Castillo, J., Estanyol, J.M., Ballescà, J.L., Oliva, R. **Proteomic characterization of the human sperm nucleus.** Proteomics 2011; 11: 2714–2726

Frapsauce, C., Pionneau, C., Bouley, J., Delarouziere, V., Berthaut, I., Ravel, C., Antoine, J.M., Soubrier, F., Mandelbaum, J. **Proteomic identification of target proteins in normal but nonfertilizing sperm.** Fertil. Steril. 2014; 102: 372–380

Freund, A., Zhong, F.L., Venteicher, A.S., Meng, Z., Veenstra, T.D., Frydman, J., Artandi, S.E. **Proteostatic control of telomerase function through TRiC-mediated folding of TCAB1.** Cell 2014; 159: 1389–1403

Fujihara, Y., Murakami, M., Inoue, N., Satouh, Y., Kaseda, K., Ikawa, M., Okabe, M. **Sperm equatorial segment protein 1, SPESP1, is required for fully fertile sperm in mouse.** J. Cell Sci. 2010; 123: 1531–1536

Goudakou, M., Kalogeraki, A., Matalliotakis, I., Panagiotidis, Y., Gullo, G., Prapas, Y. **Cryptic sperm defects may be the cause for total fertilization failure in oocyte donor cycles.** Reprod. Biomed. Online 2012; 24: 148–152

Hammoud, S.S., Nix, D.A., Zhang, H., Purwar, J., Carrell, D.T., Cairns, B.R. **Distinctive chromatin in human sperm packages genes for embryo development.** Nature 2009; 460: 473–478

Jodar, M., Barrachina, F., Oliva, R. **The use of sperm proteomics in the assisted reproduction laboratory.** Garrido N., Rivera R. A Practical Guide to Sperm Analysis CRC Press Boca Raton (FL) 2017: 233–244

Jodar, M., Selvaraju, S., Sendler, E., Diamond, M.P., Krawetz, S.A. **Reproductive Medicine Network. The presence, role and clinical use of spermatozoal RNAs.** Hum. Reprod. Update 2013; 19: 604–624

Jodar, M., Sendler, E., Moskovtsev, S.I., Librach, C.L., Goodrich, R., Swanson, S., Hauser, R., Diamond, M.P., Krawetz, S.A. **Absence of sperm RNA elements correlates with idiopathic male infertility.** Sci. Transl. Med. 2015; 7: 295re6

Krausz, C., Escamilla, A.R., Chianese, C. **Genetics of male infertility: from research to clinic.** Reproduction 2015; 150: R159–R174

Légaré, C., Droit, A., Fournier, F., Bourassa, S., Force, A., Cloutier, F., Tremblay, R., Sullivan, R. **Investigation of male infertility using quantitative comparative proteomics.** J. Proteome Res. 2014; 13: 5403–5414

Leitner, A., Joachimiak, L.A., Bracher, A., Mönkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B., Ludtke, S., Chiu, W., Hartl, F.U., Aebersold, R., Frydman, J. **The Molecular Architecture of the Eukaryotic Chaperonin TRiC/CCT.** Structure 2012; 20: 814–825

Leyens, G., Knoops, B., Donnay, I. **Expression of peroxiredoxins in bovine oocytes and embryos produced in vitro.** Mol. Reprod. Dev. 2004; 69: 243–251

Liu, X., Liu, G., Liu, J., Zhu, P., Wang, J., Wang, Y., Wang, W., Li, N., Wang, X., Zhang, C., Shen, X., Liu, F. **ITRAQ-based analysis of sperm proteome from normozoospermic men achieving the rescue-ICSI pregnancy after the IVF failure.** Clin. Proteomics 2018; 15: 27

Martínez-Heredia, J., Estanyol, J.M., Ballescà, J.L., Oliva, R. **Proteomic identification of human sperm proteins.** Proteomics 2006; 6: 4356–4369

McReynolds, S., Dzieciatkowska, M., Stevens, J., Hansen, K.C., Schoolcraft, W.B., Katz-Jaffe, M.G. **Toward the identification of a subset of unexplained infertility: A sperm proteomic approach.** Fertil. Steril. 2014; 102: 692–699

Skakkebaek, N.E., Rajpert-De Meyts, E., Buck Louis, G.M., Toppari, J., Andersson, A.-M., Eisenberg, M.L., Jensen, T.K., Jorgensen, N., Swan, S.H., Sapra, K.J., Ziebe, S., Priskorn, L., Juul, A. **Male Reproductive Disorders and Fertility Trends: Influences of Environment and Genetic Susceptibility.** Physiol. Rev. 2015; 96: 55–97

Swain, J.E., Pool, T.B. **ART failure: Oocyte contributions to unsuccessful fertilization.** Hum. Reprod. Update 2008; 14: 431–446

Torra-Massana, M., Barragán, M., Bellu, E., Oliva, R., Rodríguez, A., Vassena, R. **Sperm telomere length in donor samples is not related to ICSI outcome.** J. Assist. Reprod. Genet. 2018; 35: 649–657

Turner, S., Wong, H.P., Rai, J., Hartshorne, G.M. **Telomere lengths in human oocytes, cleavage stage embryos and blastocysts.** Mol. Hum. Reprod. 2010; 16: 685–694

Wang, G., Guo, Y., Zhou, T., Shi, X., Yu, J., Yang, Y., Wu, Y., Wang, J., Liu, M., Chen, X., Tu, W., Zeng, Y., Jiang, M., Li, S., Zhang, P., Zhou, Q., Zheng, B., Yu, C., Zhou, Z., Guo, X., Sha, J. **In-depth proteomic analysis of the human sperm reveals complex protein compositions.** J. Proteomics 2013; 79: 114–122

World Health Organization. 2010 **WHO laboratory manual for the Examination and processing of human semen.** World Health Organization

Yang, Q., Zhao, F., Dai, S., Zhang, N., Zhao, W., Bai, R., Sun, Y. **Sperm telomere length is positively associated with the quality of early embryonic development.** Hum. Reprod. 2015; 30: 1876–1881

Zhu, Y., Wu, Y., Jin, K., Lu, H., Liu, F., Guo, Y., Yan, F., Shi, W., Liu, Y., Cao, X., Hu, H., Zhu, H., Guo, X., Sha, J., Li, Z., Zhou, Z. **Differential proteomic profiling in human spermatozoa that did or did not result in pregnancy via IVF and AID.** Proteomics. Clin. Appl. 2013; 7: 850–858

Received 22 May 2019; received in revised form 31 October 2019; accepted 6 January 2020.

4.

DISCUSSIÓ

4. DISCUSSIÓ

Tradicionalment, la idea clàssica que es tenia sobre la funció de l'espermatozoide era que servia exclusivament pel transport del genoma haploide patern fins a l'oòcit madur. En part aquesta idea es derivava del fet que l'espermatozoide és una cèl·lula que no transcriu ni tradueix, degut a la pèrdua de la major part del citoplasma per la gota citoplasmàtica i al fort empaquetament que li confereixen els toroides de protamines, resultant, per tant, en una cèl·lula genèticament "inerta" (Mezquita, 1985; Cooper, 2005; Oliva, 2006). En els últims anys, gràcies als avenços en les tècniques d'alt rendiment com són la genòmica, transcriptòmica i la proteòmica, s'ha observat que la cèl•lula espermàtica madura també aporta altres components en el moment de la fertilització com són marques epigenètiques, RNAs i proteïnes, els quals podrien ser claus en els primers estadis del desenvolupament embrionari primerenc i la generació d'un nou individu sa (Jodar *et al.*, 2013; Castillo *et al.*, 2018; Jodar, 2019). Ja que l'espermatozoide madur és una cèl·lula inerta a nivell transcripcional i traduccional, el seu estudi representa un mirall de la contribució que aporta en el moment de la fertilització. Tot i així, la major part dels components moleculars de l'espermatozoide són romanents de l'espermatogènesi i per tant, reflecteix esdeveniments passats. És a dir, en la cèl·lula espermàtica madura conviuen molècules amb diferent funcionalitat, tan les molècules necessàries per la correcta fertilització i desenvolupament embrionari com romanents de l'espermatogènesi que poden mostrar defectes passats durant el procés de diferenciació cel·lular i donar explicació en alguns casos d'infertilitat (Jodar *et al.*, 2017; Castillo *et al.*, 2018).

Els estudis descriptius basats en la identificació de diferents tipus moleculars utilitzant tècniques d'alt rendiment, aporten coneixement de la composició i potencial funció específica de la cèl·lula espermàtica. D'aquesta manera, la inferència *in silico* de possibles rols i funcions de les molècules aportades ajuden a la comprensió del seu paper en la funcionalitat de l'espermatozoide i la seva contribució a l'embrió en desenvolupament. De manera complementària, els estudis comparatius en individus normozoospèrmics amb infertilitat d'origen desconegut associats a la qualitat embrionària i a l'èxit de les TRAs, permeten atribuir funcions específiques i determinar alteracions en la contribució paterna al zigot. En aquest sentit, en aquesta Tesi Doctoral, en els treballs 1 i 3 hem realitzat dues aproximacions descriptives amb inferència *in silico*, mitjançant RNA-seq i *top-down* MS, respectivament, mentre que en el treball 4 hem realitzat un estudi comparatiu mitjançant proteòmica comparativa amb marcatge diferencial. En tots ells, el propòsit ha estat augmentar el coneixement de la contribució paterna a l'embriogènesi primerenca.

4.1. Contribució del contingut de RNAs circulars (circRNAs) de l'espermatozoide a l'embrió

A nivell d'RNA, l'ús de *microarrays* i l'RNA-seq, com a eina descriptiva, han permès identificar una població complexa i específica d'RNAs tan codificants com no codificants de la cèl·lula espermàtica madura amb potencials rols més enllà de la fertilització (Ostermeier *et al.*, 2002, 2004; Krawetz *et al.*, 2011; Jodar *et al.*, 2013; Sendler *et al.*, 2013; Johnson *et al.*, 2015; Pantano *et al.*, 2015; Schuster *et al.*, 2016; Jodar, 2019). Tot i que la majoria d'RNAs codificants de l'espermatozoide madur es troben fragmentats, segurament perquè ja han realitzat la seva funció durant l'espermatogènesi, els RNAs codificants que es troben intactes es podrien traduir utilitzant la maquinària traduccional materna i presentar un paper durant les primeres etapes del desenvolupament embrionari (Ostermeier *et al.*, 2004; Krawetz, 2005; Boerke *et al.*, 2007; Jodar *et al.*, 2013, 2016a; Sendler *et al.*, 2013; Schuster *et al.*, 2016; Gòdia *et al.*, 2018; Estill *et al.*, 2019; Jodar, 2019). A més dels RNAs codificants, la cèl·lula espermàtica madura presenta una població complexa d'RNAs no codificants, que poden regular l'expressió gènica i contribuir a l'embriogènesi (Krawetz *et al.*, 2011; Jodar *et al.*, 2013; Pantano *et al.*, 2015).

Recentment, s'ha descrit un nou tipus d'RNA no codificant en altres tipus cel·lulars, els circRNAs, que regulen l'expressió gènica tan a nivell transcripcional com posttranscripcional (Salzman *et al.*, 2012; Jeck *et al.*, 2013; Memczak *et al.*, 2013; Jeck and Sharpless, 2014; Barrett and Salzman, 2016; Patop *et al.*, 2019). Els circRNAs són sintetitzats per la RNA polimerasa II i es formen per *backsplicing*. El circRNA que en resulta no té extrem 3', la qual cosa li confereix més estabilitat que les formes lineals, degut a la seva resistència a la degradació per endonucleases i exonucleases i presenten vides mitjanes més llargues que els RNAs lineals (Jeck *et al.*, 2013; Memczak *et al.*, 2013; Barrett and Salzman, 2016; Enuka *et al.*, 2016; Szabo and Salzman, 2016). Les funcions que se li han atribuït als circRNAs són la de fer "d'esponja" de miRNAs o de ribonucleoproteïnes (RBPs), reclutar proteïnes, regular *l'splicing* alternatiu i regular l'RNA polimerasa II (Hansen *et al.*, 2013; Memczak *et al.*, 2013; Zhang *et al.*, 2013, 2014; Kramer *et al.*, 2015; Starke *et al.*, 2015; Kristensen *et al.*, 2019). Recentment, alguns estudis també han proposats que certs circRNAs es podrien traduir a pèptids i podrien actuar com a proteïnes funcionals (Wang and Wang, 2015; Pamudurti *et al.*, 2017; Yang *et al.*, 2017; Kristensen *et al.*, 2019; Tang *et al.*, 2020). Tot i així, el seu potencial rol en la fertilització i el desenvolupament embrionari primerenc ha estat poc explorat (Dang *et al.*, 2016; Chioccarelli *et al.*, 2019; Ragusa *et al.*, 2019; Gòdia *et al.*, 2020; Tang *et al.*, 2020).

En el treball 1 d'aquesta Tesi Doctoral, vam descriure fins a 5830 circRNAs exònics continguts a l'espermatozoide humà madur, dels quals 737 es trobaven a tots els replicats biològics i estaven enriquits en termes d'ontologia gènica com regulació de l'expressió gènica. També, vam reportar per primera vegada que els IEs més

abundants prèviament descrits (Jodar *et al.*, 2015) i que són específics d'espermatozoide són, de fet, circRNAs intrònics. L'anàlisi de la resistència al tractament amb RNasa R va determinar que els circRNAs de l'espermatozoide, contràriament als circRNAs de les cèl·lules somàtiques en què conviuen les formes lineal i circular (Salzman *et al.*, 2012), es troben exclusivament en la seva forma circular. El fet de que les formes lineals d'aquests circRNAs continguts exclusivament en forma circular a l'espermatozoide coexistissin de forma majoritària amb la forma circular en la mostra de testicle, suggereix que les formes lineals es degraden al llarg de l'espermatogènesi i les formes circulars són retingudes selectivament a l'espermatozoide humà madur, en concordança al que s'ha descrit en ratolí (Tang *et al.*, 2020). L'estabilitat d'aquests circRNAs continguts específicament en la seva forma circular, com el circRNA derivat de l'intró 14 del gen *tripartite motif-containing protein 66* (*TRIM66),* es podrien utilitzar com a marcadors d'integritat de l'RNA de l'espermatozoide. La sensibilitat al tractament amb RNasa R d'alguns d'aquests circRNAs en una mostra d'RNA permetria determinar la baixa integritat d'aquesta, la qual cosa permetria suplir la manca d'aquest tipus de controls de qualitat en el gàmeta masculí, similar a l'ús estès del RIN en la cèl·lula somàtica (Schroeder *et al.*, 2006). Tot i així, no tots els circRNAs es mantenen intactes, sinó que alguns presenten una descircularització no aleatòria. Aquest fenomen de descircularització no estocàstica condueix a pensar que els circRNAs de la cèl·lula espermàtica podrien coexistir amb la seva maquinària de renovació, ja sigui a través de la presència de miRNAs o endoribonucleases (Hansen *et al.*, 2011; Kleaveland *et al.*, 2018; Park *et al.*, 2019) que causarien un trencament selectiu d'alguns circRNAs durant la maduració espermàtica o del trànsit fins arribar a l'oòcit, tot i que el mecanisme exacte roman desconegut.

El fet de que alguns dels 648 SREs requerits per a la concepció natural (Jodar *et al.*, 2015) corresponguin als circRNAs exònics descrits en el nostre estudi, posa de manifest la potencial associació dels circRNAs de l'espermatozoide a l'estat fèrtil de l'individu, semblant al que s'ha descrit prèviament (Chioccarelli *et al.*, 2019; Gòdia *et al.*, 2020; Tang *et al.*, 2020). De fet, els nostres resultats es troben recolzats per altres estudis que han identificat circRNAs diferencialment expressats en mostres d'elevada i baixa qualitat espermàtica o bé en el fet de que el número de circRNAs en donants fèrtils és superior al de pacients infèrtils (Chioccarelli *et al.*, 2019; Tang *et al.*, 2020). Per tant, desregulacions en el *cargo* de circRNAs de l'espermatozoide madur podrien alterar la contribució paterna al zigot i això veure's reflectit en la incapacitat d'obtenir descendència i per tant, provocar infertilitat. De manera similar, s'ha proposat que els RNAs de l'espermatozoide madur podrien trobar-se desregulats degut a estils de vida tòxics, com ara l'obesitat i alterar la seva correcta funcionalitat (Donkin *et al.*, 2016; Swanson *et al.*, 2020). En aquest sentit, alguns dels 487 SREs associats a obesitat, corresponen a alguns circRNAs (Swanson *et al.*, 2020). Aquests resultats apunten que alteracions en el perfil de circRNAs de

l'espermatozoide madur, podrien alterar el correcte funcionament de la cèl·lula espermàtica i afectar al l'embriogènesi. A més, la identificació de circRNAs continguts al zigot i a l'espermatozoide, però no a l'oòcit, suggereix que alguns circRNAs embrionaris podrien ser aportats exclusivament per la via germinal paterna (*paternally-derived* circRNAs embrionaris), la qual cosa indica la potencial necessitat d'aquests circRNAs en el correcte desenvolupament embrionari.

Tot i que la funció més popular pels circRNAs és la d'actuar com "esponges" de miRNAs (Hansen *et al.*, 2013; Barrett and Salzman, 2016; Greene *et al.*, 2017), els nostres resultats apunten que no sembla que sigui una funció principal en la cèl·lula espermàtica. Una altra funció que recentment està guanyant importància és la traducció dels circRNAs a pèptids funcionals (Wang and Wang, 2015; Legnini *et al.*, 2017; Yang *et al.*, 2017; Tang *et al.*, 2020). De fet, s'ha proposat que els circRNAs de l'espermatozoide es podrien traduir (Ragusa *et al.*, 2019) i en ratolí s'ha demostrat que la presència de *open reading frames* (ORFs) intactes, juntament amb la modificació N6-metiladenosina (m^6A) en el codó d'inici de circRNAs de les cèl·lules germinals masculines són traduïts a pèptids (Tang *et al.*, 2020). Les nostres dades recolzen aquesta hipòtesi, ja que els circRNAs validats experimentalment i els *paternally-derived* circRNAs embrionaris tenen almenys 1 ORF que permetria la traducció a pèptids funcionals. Sorprenentment, la predicció de traducció dels *paternally-derived* circRNAs embrionaris derivats dels gens *mitochondrial ribosome associated GTPase 1* (*MTG1)* i *serine/threonine-protein kinase VRK1 (VRK1)* sembla que donarien lloc a proteïnes truncades de la regió catalítica. Com que en l'espermatogènesi s'han descrit quinases truncades que poden realitzar la seva funció (Kierszenbaum, 2006), aquests circRNAs podrien ser capaços de generar proteïnes truncades funcionals. El circRNA exònic *VRK1,* el qual també hem demostrat la seva descircularització no estocàstica, podria dur a terme un potencial rol més enllà de la fertilització. La proteïna VRK1 funcional, és una quinasa que fosforila histones i estimula l'activitat telomerasa (Choi *et al.*, 2012; Yan *et al.*, 2017). Ja que s'ha demostrat la importància dels telòmers en el desenvolupament embrionari (Turner *et al.*, 2010), seria possible que el circRNA derivat del gen *VRK1* es trobés associat al manteniment de la mida dels telòmers en l'embrió en desenvolupament. Com que els circRNAs són més estables i resistents i posseeixen vides mitjanes més llargues que les seves formes lineals, els circRNAs espermàtics, com el derivat de *VRK1,* es podrien mantenir en la seva forma circular per protegir funcions essencials fins al moment de la fertilització i llavors ser traduïts en les primeres etapes del desenvolupament embrionari primerenc utilitzant la maquinària de l'oòcit (Tang *et al.*, 2020). Tot i així, estudis basats en peptidòmica buscant els pèptids potencialment traduïts dels circRNAs i estudis funcionals de circRNAs específics de l'espermatozoide permetran elucidar la seva contribució específica a l'embrió en desenvolupament. El següent pas en aquesta línia d'investigació serà realitzar estudis comparatius amb individus normozoospèrmics

amb infertilitat d'origen desconegut per tal determinar alteracions en el perfil de circRNAs que puguin estar afectant la qualitat embrionària i l'èxit o fracàs de les TRAs i, en conseqüència el desenvolupament embrionari.

4.2. Contribució del contingut de protamines i les seves modificacions posttraduccionals (PTMs) de l'espermatozoide a l'embrió

La contribució potencial epigenètica de l'espermatozoide al zigot, no sols ha estat atribuïda als RNAs de la cèl·lula espermàtica i a la metilació del DNA, sinó que també s'ha vist que l'estructura de la cromatina (distribució NH – NP en el genoma) i les PTMs de les histones hi poden jugar un paper important. Les histones paternes modificades, són aportades a l'embrió i s'ha proposat que tenen un paper en el control epigenètic de l'embriogènesi (Arpanahi *et al.*, 2009; Hammoud *et al.*, 2009; Oliva and Ballescà, 2012; Castillo *et al.*, 2014a, 2015; Vieweg *et al.*, 2015; Siklenka *et al.*, 2015; Terashima *et al.*, 2015; Luense *et al.*, 2016; Schagdarsurengin and Steger, 2016; Donkin and Barrès, 2018). Per exemple, en el model murí, s'ha descrit que alteracions en el patró de metilació de les histones dels espermatozoides, perjudica el desenvolupament i la supervivència de la descendència (Van Der Heijden *et al.*, 2008), la qual cosa indica la importància de l'epigenoma espermàtic en el desenvolupament embrionari i segurament en l'herència epigenètica (Siklenka *et al.*, 2015). Tot i així, poc s'ha estudiat el potencial rol que podrien tenir les protamines que empaqueten del 85 – 95 % del DNA patern i les seves PTMs associades en les primeres etapes del desenvolupament.

En el treball 2 d'aquesta Tesi Doctoral, primerament es va descriure el protocol minuciós per a la purificació de proteïnes en mamífers basat en mètodes descrits prèviament (de Yebra and Oliva, 1993) i, les modificacions necessàries per a dur a terme posteriors estudis per MS (Oliva *et al.*, 2008, 2009; de Mateo *et al.*, 2013; Codina *et al.*, 2015). Tanmateix, també, es va resumir breument els estudis que relacionaven les alteracions a nivell de proteïna del contingut de P1 i de la família de P2 amb diferents tipus d'infertilitat mitjançant tècniques convencionals àcid – urea PAGE (Lescoat *et al.*, 1988; Balhorn *et al.*, 1988; Bach *et al.*, 1990; de Yebra *et al.*, 1993, 1998; Khara, 1997; Bench *et al.*, 1998; Carrell and Liu, 2001; Corzett *et al.*, 2002; Mengual *et al.*, 2003; Aoki *et al.*, 2005b, 2005a, 2006b, 2006a; Torregrosa *et al.*, 2006; de Mateo *et al.*, 2007, 2009, 2011c, 2011b, 2013; Gázquez *et al.*, 2008; Hammadeh *et al.*, 2010; García-Peiró *et al.*, 2011; Simon *et al.*, 2011, 2014; Castillo *et al.*, 2011; Ribas-Maynou *et al.*, 2015). Endemés, aquestes aproximacions convencionals van relacionar alteracions en el ràtio P1/P2 i pre-P2/P2 a una baixa taxa d'èxit de les TRA (Khara, 1997; Nasr-Esfahani *et al.*, 2004b; Aoki *et al.*, 2005a, 2006b; de Mateo *et al.*, 2009; Simon *et al.*, 2011), indicant el potencial rol de les protamines a la fertilització

i el desenvolupament embrionari. Tot i així, els resultats d'aquests estudis presenten molta variabilitat i existeix discrepància entre alguns d'ells, segurament degut a les limitacions de les tècniques convencionals, les quals poden ser superades utilitzant tècniques d'alt rendiment.

Arran de la gran abundància a l'espermatozoide i les particularitats fisicoquímiques, fa que l'estudi de les protamines de l'espermatozoide sigui un bon model per a dur a terme l'optimització de la detecció de les seves potencials PTMs. Això es pot assolir a través de diferents combinacions d'estratègies proteòmiques (*top-down* i *bottom-up* MS) i en un futur poder-ho aplicar en identificar alteracions en PTMs de les protamines en diferents fenotips infèrtils, així com poder utilitzar-ho també en l'anàlisi de PTMs del conjunt de proteïnes de l'espermatozoide humà. En el treball 3, es van optimitzar les metodologies complementàries *top-down* MS i *bottom-up* MS, utilitzant les protamines com a model experimental, per identificar P1 i els membres de P2 i les seves PTMs associades en mostres independents d'individus normozoospèrmics amb la finalitat d'establir el perfil normal de proteoformes de les protamines. Per a l'estudi de les protamines, es requereix una aproximació diferent a les aproximacions clàssiques degut a les seves particulars característiques: (i) són extremadament bàsiques amb una seqüència d'aminoàcids extremadament rica en residus arginina, amb diversos residus cisteïna que poden formar ponts disulfur; (ii) són de mida molt petita (uns 50 aminoàcids); i (iii) les formes madures de P2 presenten una elevada similitud de seqüència entre elles. Si s'apliquessin només estratègies de *bottom-up* MS basades en la digestió amb tripsina dels residus aminoacídics arginina i lisina, les protamines quedarien degradades majoritàriament en mono- o di-pèptids, degut a l'elevat contingut d'arginines, per la qual cosa no seria possible la seva identificació per MS. Per tant, per superar aquests obstacles, es va modificar i optimitzar la digestió utilitzant proteïnasa K, que talla en menys llocs. A més, es va implementar la tècnica de *top-down* MS, que es basa en l'anàlisi de proteïnes intactes i ofereix informació única per sobre de les aproximacions *bottom-up* MS, ja que detecta la microheterogeneïtat de les proteoformes i dona informació de múltiples PTMs, variacions de seqüència i formes truncades que ocorren en una mateixa molècula (Smith and Kelleher, 2018). L'aplicació de les tècniques de *top-down* MS ha estat possible a que les protamines són molt petites i es poden identificar de manera "intacta" (sense digestió amb proteases) per MS. Similarment, s'ha descrit prèviament que la combinació d'ambdues metodologies és òptima per descriure PTMs en proteïnes similars a les protamines com serien les histones (Bonet-Costa *et al.*, 2012). Malgrat que alguns estudis han aplicat recentment l'ús de MS per estudiar les PTMs de les protamines de l'espermatozoide en ratolí i humà (Brunner *et al.*, 2014; Castillo *et al.*, 2015), l'ús de *pools* de mostres no dona informació sobre la variabilitat interindividual i són necessaris estudis amb mostres independents.

En aquest treball hem demostrat que la combinació de les tècniques de *top-down* i *bottom-up* MS és adequada per a l'estudi de les PTMs de les protamines i dona informació complementària amb una baixa variabilitat interindividual entre els pacients analitzats. En aquesta mateixa línia, l'ús de *top-down* MS és la tècnica que dona més informació i sembla ser idònia per establir el patró de proteoformes i PTMs i poder, en un futur, detectar alteracions (Smith *et al.*, 2013; Kelleher *et al.*, 2014; Smith and Kelleher, 2018). A través de *top-down* MS, s'han identificat proteoformes truncades prèviament descrites i d'altres de noves, tan per P1 com per pre-P2, algunes de les quals contenen residus modificats. Pel que fa a les PTMs de les protamines, mitjançant la tècnica de *top-down* MS s'ha identificat la P1 intacta, pre-P2 i les proteoformes madures de P2 i els seus patrons de fosforilació. Per la tècnica de *bottom-up* MS s'han identificat PTMs complementàries al *top-down* MS per la pre-P2. Sorprenentment, a més de la fosforilació, ambdues protamines duen una modificació de +61 Da, que podria correspondre al ió Zn^{2+}. A més, l'enfocament *bottom-up* MS ha permès la identificació de forma inequívoca la isoforma 2 de la P2, no anotada prèviament a UniProtKB (www.uniprot.org) (Bateman, 2019), que també hem validat a nivell d'RNA. Tot i que no s'ha identificat a nivell de proteïna el transcrit de la potencial isoforma 3 de la P2, podria ser que aquest producte no fos traduït a proteïna o bé que mitjançant les aproximacions utilitzades no s'hagi identificat.

En aquest treball, s'ha demostrat que la fosforilació és la PTM més identificada en l'espermatozoide humà madur, en concordança amb estudis previs (Chirat *et al.*, 1993; Pirhonen *et al.*, 1994b; Castillo *et al.*, 2015). Recentment, en model murí, s'ha descrit el rol crucial de la desfosforilació de P2 (Itoh *et al.*, 2019), senyalant una dinàmica de fosforilació – desfosforilació altament regulada. Alteracions en aquesta dinàmica podrien ser la causa subjacent d'infertilitat en diferents fenotips. De manera semblant, els nostres resultats, han demostrat que les formes natives són més abundants que les formes fosforilades. Aquest fet es recolza en la gran onada de desfosforilació que té lloc durant la maduració dels espermatozoides (Gusse *et al.*, 1986; Carrell *et al.*, 2007). Per tant, els residus que resten fosforilats a l'espermatozoide madur podrien ser romanents de l'espermatogènesi sense cap funció o bé actuar com a marques epigenètiques a l'embrió en desenvolupament. Ja que la unió de les protamines fosforilades al DNA és menys forta que les formes intactes (Björndahl and Kvist, 2010), les protamines fosforilades podrien funcionar com a marques epigenètiques i dirigir el reemplaçament de les protamines paternes per histones maternes en les primeres etapes del desenvolupament embrionari. Donant suport a aquesta hipòtesi, recentment s'ha demostrat que just després de la fertilització, les protamines que empaqueten el DNA patern són extensament re-fosforilades per la SRPK1 d'origen matern i la fosforilació marca l'inici de la reprogramació genòmica en què les protamines paternes seran reemplaçades per histones maternes (Gou *et al.*, 2020). Podria ser que les protamines paternes que

romanen fosforilades marquessin i dirigissin aquesta re-fosforilació després de la fertilització, però queda pendent aclarir aquest rol.

Mentre que les mostres analitzades no presentaven diferències mitjançant l'ús de la tècnica convencional àcid – urea PAGE, l'avaluació del ràtio pre-P2/P2 mitjançant dades derivades de *top-down* MS ha permès dividir els individus normozoospèrmics en dos subgrups. Veient la utilitat d'aquests ràtios derivats de *top-down* MS, el seu estudi en diferents fenotips infèrtils permetrà identificar-ne alteracions i la desregulació de proteoformes específiques podria alterar el desenvolupament de l'embrió. En aquesta direcció, recentment s'ha iniciat un nou estudi comparatiu per tal de determinar alteracions en el perfil de les proteoformes de protamines i les PTMs associades comparant diferents fenotips infèrtils utilitzant el *top-down* MS. Concretament, s'han utilitzat individus normozoospèrmics amb fertilitat provada, amb infertilitat d'origen desconegut, obesos amb índex de massa corporal (IMC) > 30 i d'edat avançada superior a 47 anys. Els resultats preliminars d'aquest estudi, apunten a que tot i que no hi ha diferències significatives en el ràtio P1/P2 obtingut amb dades derivades de *top-down* MS, sí que s'ha identificat un augment significatiu del ràtio pre-P2/P2 en individus normozoospèrmics obesos i d'edat avançada. Tot i així, és necessari l'ampliació de la mida mostral per a tenir resultats sòlids i conclusions fefaents i poder determinar alteracions en les proteoformes de protamines i les PTMs associades que podrien ser la causa de la infertilitat d'origen desconegut. La comparació de pacients amb infertilitat d'origen desconegut amb diferents resultats de TRAs ens ajudarà a desxifrar com les diferents proteoformes de les protamines poden afectar a la contribució paterna al zigot.

4.3. Contribució del contingut de proteïnes de l'espermatozoide a l'embrió

El contingut de proteïnes de l'espermatozoide que aporta en el moment de la fertilització a l'oòcit sembla ser que es troben associats a la qualitat embrionària després de l'ús de les TRAs amb valor diagnòstic i pronòstic i presenten un rol clau en el desenvolupament embrionari primerenc (Pixton *et al.*, 2004; Frapsauce *et al.*, 2009, 2014; Xu *et al.*, 2012; Zhu *et al.*, 2013; Azpiazu *et al.*, 2014; Légaré *et al.*, 2014; McReynolds *et al.*, 2014; Jodar *et al.*, 2018; Liu *et al.*, 2018; Wang *et al.*, 2018). En aquesta mateixa línia, gràcies a la compilació del proteoma de l'espermatozoide humà madur (6871 proteïnes) i la seva anàlisi, es va determinar que aproximadament l'11 % d'aquestes proteïnes podien presentar un paper fonamental en la fertilització i el desenvolupament embrionari i algunes d'elles ser aportades exclusivament per via paterna (Castillo *et al.*, 2018). Tot i així, la majoria dels estudis que comparen la contribució paterna a la qualitat embrionària i l'èxit de les TRAs utilitzaven *pools* de mostres, que tot i que homogeneïtza les diferents mostres, emmascara les

alteracions d'una mostra concreta que pot ser que presenti un fenotip similar a altres mostres però l'etiologia sigui diferent. També, solament un estudi ha explorat la determinació d'alteracions del contingut de les proteïnes que donen lloc a embrions d'elevada o baixa qualitat embrionària després de FIV-ICSI provinents de mostres d'individus normozoospèrmics, concretament en l'estadi de blastocist (McReynolds *et al.*, 2014). No obstant, es creu que l'espermatozoide té un paper fonamental durant els primers estadis del desenvolupament embrionari, abans de l'activació del genoma que té lloc entre els dies 2 i 3 entre l'estadi de 4 – 8 cèl·lules (Niakan *et al.*, 2012). A més, l'ús de mostres provinents de FIV-ICSI, permet descartar alteracions espermàtiques relacionades amb el procés de fertilització, ja que aquesta tècnica se salta el reconeixement i fusió de la cèl·lula espermàtica amb l'oòcit (Flaherty *et al.*, 1998). Tot i així, com s'ha demostrat en l'annex 3, l'ús de FIV-ICSI també pot donar explicació en alguns casos de fallida de fertilització on l'espermatozoide és incapaç d'activar l'oòcit. Els resultats de l'annex 3 suggereixen l'existència d'un procés multifactorial amb mecanismes addicionals en l'activació oocitària a més dels prèviament descrits, en el qual alteracions específiques ens proteïnes mitocondrials podrien ser la causa subjacent d'aquestes fallides.

A part dels estudis descriptius amb inferència *in silico,* la comparació d'espermatozoides provinents d'individus normozoospèrmics amb la taxa d'èxit de diferents tipus de TRAs sense un factor femení conegut, permet caracteritzar i diferenciar la contribució espermàtica en la fertilització i les diferents etapes del desenvolupament embrionari. Seguint aquesta estratègia, en el treball 4 de la present Tesi Doctoral, es va realitzar un estudi comparatiu on es correlacionava l'abundància de proteïnes de l'espermatozoide amb la qualitat d'embrions en dia 2 derivats de FIV-ICSI de pacients normozoospèrmics amb infertilitat d'origen desconegut a través de proteòmica comparativa amb marcatge diferencial. Per evitar la detecció de falsos positius, es va utilitzar la tècnica d'identificació i quantificació descrita a l'article de Barrachina i col·legues (Barrachina *et al.*, 2019) que es basa en uns criteris molt estrictes. A propòsit d'això, el fet de que hi hagi una elevada coincidència entre els resultats obtinguts en aquest treball i en estudis similars, implica que uns criteris d'identificació i quantificació estrictes permeten la disminució de la identificació de falsos positius. Un percentatge elevat de les proteïnes correlacionades amb la baixa qualitat embrionària estaven associades a l'alteració del *chaperonin TCP-1 ring complex* (TRiC). Aquest complex presenta un paper important en el plegament de moltes proteïnes, algunes d'elles relacionades amb el plegament de proteïnes associades als telòmers. Per tant, aquests resultats recolzen la hipòtesi de l'existència d'una correlació positiva entre la llargada dels telòmers i la qualitat embrionària (Turner *et al.*, 2010). A més, alguna de les proteïnes correlacionada amb la qualitat embrionària podria ser un potencial biomarcador de predicció del resultat de TRAs en parelles que presenten infertilitat d'origen desconegut, com la proteïna *uncharacterized protein C7orf61* (C7ORF61), la qual es

correlaciona amb la qualitat embrionària independentment de si es presenten alteracions en els paràmetres seminals o no.

L'anàlisi a nivell de pèptids va revelar una elevada heterogeneïtat entre els pèptids que conformen una proteïna i la seva correlació amb la qualitat embrionària. La identificació i caracterització de pèptids i proteïnes en MS es basa en l'obtenció de la massa intacta de la molècula. Per tant, qualsevol modificació de la forma nativa del pèptid, per exemple una variació no-sinònima o la presència de PTMs, portarà com a conseqüència que el pèptid intacte no serà identificat i quantificat. D'aquesta manera, l'elevada heterogeneïtat podria ser degut a la presencia de PTMs o isoformes. Tot i així, els resultats d'aquest estudi donen suport a que les proteïnes de l'espermatozoide presenten un paper en el desenvolupament embrionari primerenc i poden afectar la qualitat embrionària després de l'ús de les TRAs.

4.4. Contribució de l'espermatozoide més enllà del desenvolupament embrionari: Herència epigenètica

En aquesta Tesi Doctoral, s'ha demostrat la contribució paterna a l'embrió en desenvolupament a diferents nivells moleculars, ja siguin els circRNAs, les PTMs de les protamines o proteïnes que afecten la qualitat embrionària. Però més enllà del seu paper en la fertilitat i el desenvolupament embrionari primerenc a través de la regulació de l'expressió gènica, recentment s'ha posat de manifest el rol que alguns d'aquests factors moleculars de l'espermatozoide podrien participar en l'herència epigenètica de fenotips associats a les exposicions ambientals i estil de vida del pare (Rassoulzadegan *et al.*, 2006; Rando, 2012; Dias and Ressler, 2014; Sharma and Rando, 2014; Gapp *et al.*, 2014; Rodgers *et al.*, 2015; Terashima *et al.*, 2015; Grandjean *et al.*, 2015; Chen *et al.*, 2016b; Schagdarsurengin and Steger, 2016; Sharma *et al.*, 2016; Xavier *et al.*, 2019; Zhang *et al.*, 2019b; Jodar, 2019; Perez and Lehner, 2019; Galan *et al.*, 2020; Gou *et al.*, 2020).

Per exemple, s'ha vist en espècies model que la injecció de certs subtipus d'RNAs o RNAs totals d'espermatozoides de mascles exposats a estímuls ambientals, com dieta o estressants, en zigots sans, indueix que la progènie presenti totalment o en part les alteracions fenotípiques que presentava la descendència natural dels mascles exposats, incloent canvis metabòlics o de comportament (Xavier *et al.*, 2019; Galan *et al.*, 2020). En humà, de forma interessant, alteracions en el *cargo* d'RNAs han estat associats a trobar-se desregulats degut a l'estil de vida, com ara en l'obesitat, alterant la correcta funció espermàtica (Donkin *et al.*, 2016; Swanson *et al.*, 2020). Recentment, s'han identificat SREs presents com a transcrit complet, en l'espermatozoide en la comparació d'individus obesos amb individus prims (Swanson *et al.*, 2020). Concretament, s'han identificat 16 cofactors de la remodelació de la cromatina, 11 positivament i 5 negativament correlacionats amb

l'índex de massa corporal (Swanson *et al.*, 2020). Aquests SREs presents com a transcrit complet podrien jugar un paper en el correcte funcionament de l'espermatozoide i alteracions del perfil d'RNAs lligades a l'obesitat, podrien afectar la funció espermàtica, tenir un paper fonamental durant les primeres etapes de l'embriogènesi i potser modular l'herència epigenètica.

Tot i el potencial rol que poden jugar els RNAs codificants a l'herència epigenètica, l'estudi dels RNAs no codificants ha estat clau per identificar el potencial rol de l'espermatozoide en la salut de la descendència. Degut a les limitacions ètiques que presenta el model humà, la inferència de l'herència epigenètica per la via germinal masculina per part dels RNAs no codificants ha estat feta a través de models animals, principalment murí. La injecció de 9 miRNAs abundants de l'espermatozoide de mascles que patien estrès crònic en zigots sans, provocava la degradació de mRNAs materns, alguns d'ells implicats en la remodelació de la cromatina i la seva descendència patia desordres en el seu comportament associats a l'estrès (Rodgers *et al.*, 2015). Un altre estudi va demostrar que dependent de si s'injectaven *small* o *long* RNAs de ratolins amb estrès traumàtic a zigots sans, la descendència adquiria fenotips desiguals, suggerint rols diferencials segons els tipus d'RNAs involucrats (Gapp *et al.*, 2014). Un dels models més empleats és l'ús de rosegadors sotmesos a dietes poc sanes, ja siguin altes en greixos (*high-fat diet*; HFD) o baixes en proteïnes, per tal de determinar l'herència epigenètica dels fenotips alterats associats a l'estil de vida del pare. En un estudi, la injecció del miR-19b provinent d'espermatozoides d'individus sota una dieta HFD a zigots sans, provocava que la descendència fos obesa i tingués intolerància a la glucosa, la qual durava vàries generacions (Grandjean *et al.*, 2015). També es va demostrar que una dieta baixa en proteïnes altera el perfil de miRNAs, piRNAs i tRFs (Sharma *et al.*, 2016). Els tRFs són porcions provinents dels RNAs transferència (tRNAs) madurs i es troben enriquits a l'espermatozoide (Peng *et al.*, 2012; Gòdia *et al.*, 2018). Solen contenir l'extrem 5' entre el D-*loop* i el *loop* anticodó, però en espermatozoide també s'han trobat fragments que contenen l'extrem 3' (Jodar *et al.*, 2013). En ratolí, s'ha vist que l'espermatozoide augmenta el seu número al llarg de la maduració epididimària (Sharma *et al.*, 2016) i s'ha proposat que presenten un paper clau en l'herència epigenètica de l'obesitat associada a la dieta paterna (Chen *et al.*, 2016a; Zhang *et al.*, 2019b). La injecció de tRFs provinents d'espermatozoides de ratolins HFD en zigots WT provocava alteracions metabòliques a la descendència, tot i que el fenotip no era igual que la descendència natural del pares sotmesos a dieta HFD (Chen *et al.*, 2016a). Així doncs, ha estat validat el paper dels RNAs no codificants en l'herència epigenètica de fenotips alterats associats a les experiències paternes com dietes poc saludables, traumes o estrès (Rassoulzadegan *et al.*, 2006; Rando, 2012; Grandjean *et al.*, 2015; Chen *et al.*, 2016b; Schagdarsurengin and Steger, 2016; Sharma *et al.*, 2016; Zhang *et al.*, 2019b).

Però no sols les marques epigenètiques del DNA i els RNAs semblen estar involucrats en l'herència epigenètica per la via paterna, ja que proteïnes amb o sense PTMs de l'espermatozoide amb funció epigenètica podrien modular l'expressió gènica de l'embrió i modular el fenotip de la progènie. Recentment, en l'estudi *in silico* de Castillo i col·legues es van identificar 3 proteïnes (*Eukaryotic Translation Elongation Factor 1 Alpha 1* (EEF1A), *CCR4-NOT Transcription Complex Subunit 1* (CNOT1) i *Elongator Acetyltransferase Complex Subunit 3* (ELP3)) implicades en l'expressió gènica que estaven desregulades en individus obesos (Castillo *et al.*, 2018). A tall d'exemple, la proteïna ELP3 es troba relacionada amb la modificació de tRNAs i és una de les molècules encarregades de la desmetilació del DNA patern després de la fertilització (Svejstrup, 2007; Okada *et al.*, 2010; Castillo *et al.*, 2018). En un altre estudi recent, es van comparar les proteïnes de l'espermatozoide d'individus obesos i individus prims. La comparació va resultar en la identificació en abundància disminuïda en els individus obesos de les proteïnes *eukaryotic translation initiation factor 3 subunit F* (EIF3F) i *eukaryotic initiation factor 4A-II* (EIF4A2) que es troben involucrades en la iniciació de la traducció. EIF3F inhibeix la síntesi de proteïnes a nivell traduccional, mentre que EIF4A2 es troba implicat en el silenciament de mRNAs a través de miRNAs (Pini *et al.*, 2020). Aquestes troballes, reforcen la idea que les proteïnes també podrien tenir un paper rellevant tan en el desenvolupament embrionari primerenc com en l'herència epigenètica. En el cas particular de les protamines, recentment s'ha demostrat que la fosforilació de les protamines marca l'inici de la reprogramació genòmica en què les protamines paternes seran reemplaçades per histones maternes (Gou *et al.*, 2020). Els resultats preliminars de l'estudi comparatiu de les PTMs de les protamines humanes, han demostrat que individus normozoospèrmics obesos i d'edat avançada presenten ràtios de pre-P2/P2 derivats de *top-down* MS significativament diferents als obtinguts en el grup control (dades no mostrades), donant suport a la hipòtesi que alteracions en el perfil de PTMs de les protamines, podrien aportar a l'embrió un perfil alterat, donar lloc a una reprogramació genòmica aberrant i afectar la salut de la descendència, tot i que és necessària l'ampliació del número de mostres.

Malgrat tots els estudis que recolzen que les exposicions ambientals i estil de vida del pare pot afectar la salut de la descendència, la inferència de l'herència epigenètica al llarg de diferents generacions en humà, només s'ha pogut realitzar, per motius obvis, mitjançant estudis retrospectius, que han validat, també, la transmissió d'adaptacions ambientals per la línia germinal paterna, però en els quals no es poden determinar les molècules específiques implicades. Estudiant èpoques de fam i abundància, es va determinar que una sobrealimentació durant la infantesa de l'home, provocava que els seus fills i néts tinguessin una esperança de vida inferior degut a l'augment de risc de contraure diabetis. Per contra, la fam durant la infantesa de l'home, disminuïa en els seus fills i néts el risc de mort per causes cardiovasculars i diabetis (Bygren *et al.*, 2001; Kaati *et al.*, 2002). En un estudi més recent, es va

concloure que una sobrealimentació paterna a la infantesa provocava als néts un increment en el risc de morir de càncer, però no a les nétes (Vågerö *et al.*, 2018).

En conclusió, el conjunt de resultats dels diferents treballs d'aquesta Tesi Doctoral, han permès millorar la comprensió de la contribució paterna al desenvolupament embrionari primerenc. Els estudis descriptius han permès determinar que en el moment de la fertilització, l'espermatozoide aporta a l'oòcit una població complexa i única de circRNAs alguns d'ells aportats exclusivament pel llinatge patern, així com el DNA empaquetat amb protamines que contenen PTMs associades amb potencials rols epigenètics. L'estudi comparatiu ha contribuït a caracteritzar un conjunt de proteïnes crucials relacionades amb la qualitat embrionària i per tant, en l'embriogènesi. Nous estudis comparant tan la població de circRNAs com les proteoformes de les protamines i les seves PTMs associades amb diferents fenotips, permetrà augmentar la comprensió sobre la seva contribució a la fertilització, el desenvolupament embrionari i més enllà.

5.

CONCLUSIONS

5. CONCLUSIONS

Les conclusions derivades dels resultats obtinguts d'aquesta Tesi Doctoral són les següents:

1. L'espermatozoide conté una població abundant i complexa d'RNAs circulars (circRNAs) tan exònics com intrònics. De fet, els elements intrònics (IEs) més abundants i específics d'espermatozoide detectats en estudis previs són circRNAs intrònics.

2. A diferència dels circRNA identificats en les cèl·lules somàtiques, els circRNA de l'espermatozoide s'han detectat principalment en la seva forma circular, sense presència dels seus equivalents lineals, que són degradats durant l'espermatogènesi.

3. L'estabilitat observada entre mostres dels circRNAs de l'espermatozoide continguts únicament en la seva forma circular, com per exemple el circRNA derivat de l'intró 14 del gen *tripartite motif-containing protein 66 (TRIM66)*, suggereix que la resistència d'aquests circRNAs al tractament amb RNasa R podria ser utilitzat com un test d'integritat de l'RNA d'espermatozoide, suplint la manca de funcionalitat que presenta l'*RNA integrity number* (RIN).

4. La descircularització no estocàstica d'alguns circRNAs de l'espermatozoide, suggereix que aquests circRNAs podrien coexistir amb la seva maquinària de renovació a la cèl·lula espermàtica madura.

5. L'enriquiment en funcions de regulació de l'expressió gènica dels circRNAs detectats en totes les mostres analitzades, sumant-se a que certs circRNAs es troben associats a *sperm RNA elements* (SREs) requerits per aconseguir un naixement i/o alterats en pacients obesos i que alguns circRNAs del zigot són aportats exclusivament per via paterna, suggereix que els circRNAs retinguts específicament a l'espermatozoide poden estar implicats en la fertilitat masculina, així com en la fertilització i el desenvolupament embrionari primerenc.

6. La presència de *open reading frames* (ORFs) intactes en els circRNAs de l'espermatozoide reforça la hipòtesi que els circRNAs del gàmeta masculí podrien ser traduïts a pèptids funcionals per la maquinària de l'oòcit i presentar rols putatius a l'embriogènesi.

7. El protocol detallat d'extracció de protamines ha permès posar en coneixement de la societat científica els passos crítics d'aquest procediment. També, hem descrit les modificacions necessàries per a realitzar estudis d'espectrometria de masses (MS) posteriors a l'extracció.

8. S'ha demostrat que la combinació de les tècniques de *top-down* i *bottom-up* MS és adient per estudiar el perfil de proteoformes de les protamines humanes, sobretot el *top-down* MS, ja que proporcionen informació complementària entre elles i presenten una baixa variabilitat interindividual.

9. Mitjançant *top-down* MS s'ha establert el patró normal de fosforilació de les protamines humanes en individus normozoospèrmics, així com de les modificacions posttraduccionals (PTMs) addicionals.

10. L'aproximació per *bottom-up* MS ha permès la validació inequívoca de la isoforma 2 de P2 a nivell proteic, fins ara no anotada a UniProtKB.

11. L'estudi per *top-down* MS ha permès també detectar per primera vegada una addició de 61 Da en algunes proteoformes, indicant que aquesta modificació podria ser la unió de Zn^{2+}. Endemés, s'han identificat noves formes truncades tan per protamina 1 (P1) com pel precursor de la protamina 2 (pre-P2).

12. L'establiment dels ràtios de P1/P2 i pre-P2/P2 derivats de dades de *top-down* MS ha permès superar les limitacions que presenta l'ús de tècniques convencional en els càlculs d'aquests ràtios. S'ha demostrat que mentre que el ràtio P1/P2 derivat de dades de *top-down* MS roman estable entre els pacients analitzats, el ràtio pre-P2/P2 permet estratificar els individus normozoospèrmics en subgrups, cosa que no permetia l'ús de tècniques convencionals.

13. La correlació de l'abundància de proteïnes de l'espermatozoide amb la qualitat d'embrions derivats de fecundació *in vitro* amb injecció intracitoplasmàtica (FIV-ICSI) en dia 2 del desenvolupament embrionari provinents de pacients sense cap factor masculí o femení conegut, contribueix a aprofundir la implicació de l'espermatozoide en l'embriogènesi primerenca.

14. La proteïna *uncharacterized protein C7orf61* (C7ORF61) de l'espermatozoide podria ser utilitzada com a biomarcador de la predicció de la taxa d'èxit de diferents tècniques de reproducció assistida, ja que es correlaciona amb la qualitat embrionària independentment de si es presenten alteracions en els paràmetres seminals.

15. L'aplicació de criteris estrictes d'identificació i quantificació de proteïnes quan s'utilitzen tècniques d'alt rendiment redueix el nombre de falsos positius, derivant en resultats més robusts i d'elevada coincidència amb estudis similars.

16. L'anàlisi de les dades proteòmiques a nivell de pèptids ha revelat una elevada heterogeneïtat entre els pèptids que conformen una proteïna i la seva correlació amb la qualitat embrionària. Es requereixen, doncs, noves estratègies per complementar les dades proteòmiques amb informació de PTMs i isoformes.

17. La combinació de dos estudis descriptius amb inferència *in silico* i un estudi comparatiu utilitzant tècniques d'alt rendiment ha donat informació complementària per determinar l'impacte de la cèl·lula espermàtica al desenvolupament embrionari primerenc i a la generació d'un nou individu sa.

6.

REFERÈNCIES

6. REFERÈNCIES

Abu-Halima M, Hammadeh M, Schmitt J, Leidinger P, Keller A, Meese E, Backes C. Altered microRNA expression profiles of human spermatozoa in patients with different spermatogenic impairments. *Fertil Steril* 2013;**99**:1249-1255.e16.

Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003;**422**:198–207.

Aitken RJ, Nixon B. Sperm capacitation: a distant landscape glimpsed but unexplored. *Mol Hum Reprod* 2013;**19**:785–793.

Amaral A, Castillo J, Estanyol JM, Ballescà JL, Ramalho-Santos J, Oliva R. Human Sperm Tail Proteome Suggests New Endogenous Metabolic Pathways. *Mol Cell Proteomics* 2013;**12**:330–342.

Amaral A, Castillo J, Ramalho-Santos J, Oliva R. The combined human sperm proteome: Cellular pathways and implications for basic and clinical science. *Hum Reprod Update* 2014a;**20**:40–62.

Amaral A, Paiva C, Attardo Parrinello C, Estanyol JM, Ballescà JL, Ramalho-Santos J, Oliva R. Identification of proteins involved in human sperm motility using high-throughput differential proteomics. *J Proteome Res* 2014b;**13**:5670–5684.

Aoki VW, Emery BR, Liu L, Carrell DT. Protamine levels vary between individual sperm cells of infertile human males and correlate with viability and DNA integrity. *J Androl* 2006a;**27**:890–898.

Aoki VW, Liu L, Carrell DT. Identification and evaluation of a novel sperm protamine abnormality in a population of infertile males. *Hum Reprod* 2005a;**20**:1298–1306.

Aoki VW, Liu L, Jones KP, Hatasaka HH, Gibson M, Peterson CM, Carrell DT. Sperm protamine 1/protamine 2 ratios are related to in vitro fertilization pregnancy rates and predictive of fertilization ability. *Fertil Steril* 2006b;**86**:1408–1415.

Aoki VW, Moskovtsev SI, Willis J, Liu L, Mullen JBM, Carrell DT. DNA integrity is compromised in protamine-deficient human sperm. *J Androl* 2005b;**26**:741–748.

Arato I, Grande G, Barrachina F, Bellucci C, Lilli C, Jodar M, Aglietti MC, Mancini F, Vincenzoni F, Pontecorvi A, *et al.* "In vitro" Effect of Different Follicle—Stimulating Hormone Preparations on Sertoli Cells: Toward a Personalized Treatment for Male Infertility. *Front Endocrinol (Lausanne)* 2020;**11**:.

Arpanahi A, Brinkworth M, Iles D, Krawetz SA, Paradowska A, Platts AE, Saida M, Steger K, Tedder P, Miller D. Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding

sequences. *Genome Res* 2009;**19**:1338–1349.

Auclair G, Weber M. Mechanisms of DNA methylation and demethylation in mammals. *Biochimie* 2012;**94**:2202–2211.

Ausió J. The shades of gray of the chromatin fiber. *BioEssays* 2015;**37**:46–51.

Avenarius MR, Hildebrand MS, Zhang Y, Meyer NC, Smith LLH, Kahrizi K, Najmabadi H, Smith RJH. Human Male Infertility Caused by Mutations in the CATSPER1 Channel Protein. *Am J Hum Genet* 2009;**84**:505–510.

Avidor-Reiss T, Fishman EL. It takes two (centrioles) to tango. *Reproduction* 2019;**157**:R33–R51.

Aydin H, Sultana A, Li S, Thavalingam A, Lee JE. Molecular architecture of the human sperm IZUMO1 and egg JUNO fertilization complex. *Nature* 2016;**534**:562–565.

Azpiazu R, Amaral A, Castillo J, Estanyol JM, Guimerà M, Ballescà JL, Balasch J, Oliva R. High-throughput sperm differential proteomics suggests that epigenetic alterations contribute to failed assisted reproduction. *Hum Reprod* 2014;**29**:1225–1237.

Bach O, Glander HJ, Scholz G, Schwarz J. Electrophoretic patterns of spermatozoal nucleoproteins (NP) in fertile men and infertility patients and comparison with NP of somatic cells. *Andrologia* 1990;**22**:217–224.

Baker MA, Hetherington L, Weinberg A, Naumovski N, Velkov T, Pelzing M, Dolman S, Condina MR, Aitken RJ. Analysis of phosphopeptide changes as spermatozoa acquire functional competence in the epididymis demonstrates changes in the post-translational modification of izumo1. *J Proteome Res* 2012;**11**:5252–5264.

Baker MA, Naumovski N, Hetherington L, Weinberg A, Velkov T, Aitken RJ. Head and flagella subcompartmental proteomic analysis of human spermatozoa. *Proteomics* 2013;**13**:61–74.

Baker MA, Reeves G, Hetherington L, Müller J, Baur I, Aitken RJ. Identification of gene products present in Triton X-100 soluble and insoluble fractions of human spermatozoa lysates using LC-MS/MS analysis. *Proteomics - Clin Appl* 2007;**1**:524–532.

Balhorn R. The protamine family of sperm nuclear proteins. *Genome Biol* 2007;**8**:227.

Balhorn R, Corzett M, Mazrimas J, Stanker LH, Wyrobek A. High-performance liquid chromatographic separation and partial characterization of human protamines 1, 2, and 3. *Biotechnol Appl Biochem* 1987;**9**:82–88.

Balhorn R, Corzett M, Mazrimas JA. Formation of intraprotamine disulfides in vitro.

Arch Biochem Biophys 1992;**296**:384–393.

Balhorn R, Reed S, Tanphaichitr N. Aberrant protamine 1/protamine 2 ratios in sperm of infertile human males. *Experientia* 1988;**44**:52–55.

Ballescà JL, Balasch J, Calafell JM, Alvarez R, Fábregues F, Martinez De Osaba MJ, Ascaso C, Vanrell JA. Serum inhibin B determination is predictive of successful testicular sperm extraction in men with non-obstructive azoospermia. *Hum Reprod* 2000;**15**:1734–1738.

Ballescà JL, Oliva R. *Técnicas de reproducción asistida en el factor masculino*. In Cruz N, editor. 2012;

Bangs F, Antonio N, Thongnuek P, Welten M, Davey MG, Briscoe J, Tickle C. Generation of mice with functional inactivation of talpid3, a gene first identified in chicken. *Development* 2011;**138**:3261–3272.

Bansal SK, Gupta N, Sankhwar SN, Rajender S. Differential genes expression between fertile and infertile spermatozoa revealed by transcriptome analysis. *PLoS One* 2015;**10**:e0127007.

Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: A critical review. *Anal Bioanal Chem* 2007;**389**:1017–1031.

Barbaux S, Ialy-Radio C, Chalbi M, Dybal E, Homps-Legrand M, Do Cruzeiro M, Vaiman D, Wolf JP, Ziyyat A. Sperm SPACA6 protein is required for mammalian Sperm-Egg Adhesion/Fusion. *Sci Rep* 2020;**10**:1–15.

Barrachina F, Jodar M, Delgado-Dueñas D, Soler-Ventura A, Estanyol JM, Mallofré C, Ballescà JL, Oliva R. Stable-protein Pair Analysis as A Novel Strategy to Identify Proteomic Signatures: Application To Seminal Plasma From Infertile Patients. *Mol Cell Proteomics* 2019;**18**:S77–S90.

Barrachina F, Soler-Ventura A, Oliva R, Jodar M. Sperm Nucleoproteins (Histones and Protamines). *A Clin Guid to Sperm DNA Chromatin Damage* 2018;, p. 31–51.

Barrett SP, Salzman J. Circular RNAs: analysis, expression and potential functions. *Development* 2016;**143**:1838–1847.

Bateman A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res* 2019;**47**:D506–D515.

Bedford JM. Enigmas of mammalian gamete form and function. *Biol Rev Camb Philos Soc* 2004;**79**:429–460.

Bench G, Corzett MH, de Yebra L, Oliva R, Balhorn R. Protein and DNA contents in sperm from an infertile human male possessing protamine defects that vary over time. *Mol Reprod Dev* 1998;**50**:345–353.

Bench G, Corzett MH, Kramer CE, Grant PG, Balhorn R. Zinc is sufficiently abundant within mammalian sperm nuclei to bind stoichiometrically with protamine 2. *Mol Reprod Dev* 2000;**56**:512–519.

Beraldi R, Pittoggi C, Sciamanna I, Mattei E, Spadafora C. Expression of LINE-1 retroposons is essential for murine preimplantation development. *Mol Reprod Dev* 2006;**73**:279–287.

Bergh PA, Navot D. The impact of embryonic development and endometrial maturity on the timing of implantation. *Fertil Steril* 1992;**58**:537–542.

Bestor TH. The DNA methyltransferases of mammals. *Hum Mol Genet* 2000;**9**:2395–2402.

Bianchi E, Doe B, Goulding D, Wright GJ. Juno is the egg Izumo receptor and is essential for mammalian fertilization. *Nature* 2014;**508**:483–487.

Björkgren I, Sipilä P. The impact of epididymal proteins on sperm function. *Reproduction* 2019;**158**:R155–R167.

Björndahl L, Kvist U. Human sperm chromatin stabilization: A proposed model including zinc bridges. *Mol Hum Reprod* 2010;**16**:23–29.

Boerke A, Dieleman SJ, Gadella BM. A possible role for sperm RNA in early embryo development. *Theriogenology* 2007;**68**:147–155.

Bonache S, Mata A, Ramos MD, Bassas L, Larriba S. Sperm gene expression profile is related to pregnancy rate after insemination and is predictive of low fecundity in normozoospermic men. *Hum Reprod* 2012;**27**:1556–1567.

Bonduelle M, Wennerholm UB, Loft A, Tarlatzis BC, Peters C, Henriet S, Mau C, Victorin-Cederquist A, Van Steirteghem A, Balaska A, *et al.* A multi-centre cohort study of the physical health of 5-year-old children conceived after intracytoplasmic sperm injection, in vitro fertilization and natural conception. *Hum Reprod* 2005;**20**:413–419.

Bonet-Costa C, Vilaseca M, Diema C, Vujatovic O, Vaquero A, Omeñaca N, Castejón L, Bernués J, Giralt E, Azorín F. Combined bottom-up and top-down mass spectrometry analyses of the pattern of post-translational modifications of Drosophila melanogaster linker histone H1. *J Proteomics* 2012;**75**:4124–4138.

Botta T, Blescia S, Martínez-Heredia J, Lafuente R, Brassesco M, Luis Ballescà J, Oliva R. Identificación de diferencias proteómicas en muestras oligozoospérmicas. *Rev Int Androl* 2009;**7**:14–19.

Bourc'his D, Voinnet O. A small-RNA perspective on gametogenesis, fertilization, and early zygotic development. *Science (80-)* 2010;**330**:617–622.

Bracke A, Peeters K, Punjabi U, Hoogewijs D, Dewilde S. A search for molecular

mechanisms underlying male idiopathic infertility. *Reprod Biomed Online* 2018;**36**:327–339.

Breton S, Ruan YC, Park YJ, Kim B. Regulation of epithelial function, differentiation, and remodeling in the epididymis. *Asian J Androl* 2016;**18**:3–9.

Brucker C, Lipford GB. The human sperm acrosome reaction: Physiology and regulatory mechanisms. an update. *Hum Reprod Update* 1995;**1**:51–62.

Brunner AM, Nanni P, Mansuy IM. Epigenetic marking of sperm by post-translational modification of histones and protamines. *Epigenetics and Chromatin* 2014;**7**:1–12.

Burl RB, Clough S, Sendler E, Estill M, Krawetz SA. Sperm RNA elements as markers of health. *Syst Biol Reprod Med* 2018;**64**:25–38.

Busso D, Cohen DJ, Maldera JA, Dematteis A, Cuasnicu PS. A novel function for CRISP1 in rodent fertilization: Involvement in sperm-zona pellucida interaction. *Biol Reprod* 2007;**77**:848–854.

Bustin S. Molecular Biology of the Cell, Sixth Edition; ISBN: 9780815344643; and Molecular Biology of the Cell, Sixth Edition, The Problems Book; ISBN 9780815344537. *Int J Mol Sci* 2015;**16**:28123–28125.

Bygren LO, Kaati G, Edvinsson S. Longevity determined by paternal ancestors' nutrition during their slow growth period. *Acta Biotheor* 2001;**49**:53–59.

Caballero-Campo P, Lira-Albarran S, Barrera D, Borja-Cacho E, Godoy-Morales H, Rangel-Escareno C, Larrea F, Chirinos M. Gene transcription profiling of astheno- And normo-zoospermic sperm subpopulations. *Asian J Androl* 2020;**22**:608–615.

Cahill DJ, Wardle PG. Management of infertility. *Br Med J* 2002;**325**:28–32.

Carapito C, Duek P, Macron C, Seffals M, Rondel K, Delalande F, Lindskog C, Fréour T, Vandenbrouck Y, Lane L, *et al.* Validating Missing Proteins in Human Sperm Cells by Targeted Mass-Spectrometry- and Antibody-based Methods. *J Proteome Res* 2017;**16**:4340–4351.

Carrell DT, Aston KI, Oliva R, Emery BR, de Jonge CJ. The "omics" of human male infertility: integrating big data in a systems biology approach. *Cell Tissue Res* 2016;**363**:295–312.

Carrell DT, Emery BR, Hammoud S. Altered protamine expression and diminished spermatogenesis: what is the link? *Hum Reprod Update* 2007;**13**:313–327.

Carrell DT, Hammoud SS. The human sperm epigenome and its potential role in embryonic development. *Mol Hum Reprod* 2009;**16**:37–47.

Carrell DT, Liu L. Altered protamine 2 expression is uncommon in donors of known

fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. *J Androl* 2001;**22**:604–610.

Castillo J, Amaral A, Azpiazu R, Vavouri T, Estanyol JM, Ballesca JL, Oliva R. Genomic and proteomic dissection and characterization of the human sperm chromatin. *Mol Hum Reprod* 2014a;**20**:1041–1053.

Castillo J, Amaral A, Oliva R. Sperm nuclear proteome and its epigenetic potential. *Andrology* 2014b;**2**:326–338.

Castillo J, Bogle OA, Jodar M, Torabi F, Delgado-Dueñas D, Estanyol JM, Ballescà JL, Miller D, Oliva R. Proteomic Changes in Human Sperm During Sequential in vitro Capacitation and Acrosome Reaction. *Front cell Dev Biol* 2019;**7**:295.

Castillo J, Estanyol JM, Ballescà JL, Oliva R. Human sperm chromatin epigenetic potential: Genomics, proteomics, and male infertility. *Asian J Androl* 2015;**17**:601–609.

Castillo J, Jodar M, Oliva R. The contribution of human sperm proteins to the development and epigenome of the preimplantation embryo. *Hum Reprod Update* 2018;**24**:535–555.

Castillo J, Simon L, de Mateo S, Lewis S, Oliva R. Protamine/DNA ratios and DNA damage in native and density gradient centrifuged sperm from infertile patients. *J Androl* 2011;**32**:324–332.

Chakravarty S, Kadunganattil S, Bansal P, Sharma RK, Gupta SK. Relevance of glycosylation of human zona pellucida glycoproteins for their binding to capacitated human spermatozoa and subsequent induction of acrosomal exocytosis. *Mol Reprod Dev* 2008;**75**:75–88.

Chalbi M, Barraud-Lange V, Ravaux B, Howan K, Rodriguez N, Soule P, Ndzoudi A, Boucheix C, Rubinstein E, Wolf JP, *et al.* Binding of sperm protein Izumo1 and its egg receptor juno drives cd9 accumulation in the intercellular contact area prior to fusion during mammalian fertilization. *Dev* 2014;**141**:3732–3739.

Chan CC, Shui HA, Wu CH, Wang CY, Sun GH, Chen HM, Wu GJ. Motility and protein phosphorylation in healthy and asthenozoospermic sperm. *J Proteome Res* 2009;**8**:5382–5386.

Chao HCA, Chung CL, Pan HA, Liao PC, Kuo PL, Hsu CC. Protein tyrosine phosphatase non-receptor type 14 is a novel sperm-motility biomarker. *J Assist Reprod Genet* 2011;**28**:851–861.

Chen Q, Yan M, Cao Z, Li X, Zhang Y, Shi J, Feng GH, Peng H, Zhang X, Zhang Y, *et al.* Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* 2016a;**351**:397–400.

Chen Q, Yan W, Duan E. Epigenetic inheritance of acquired traits through sperm

RNAs and sperm RNA modifications. *Nat Rev Genet* 2016b;**17**:733–743.

Chioccarelli T, Manfrevola F, Ferraro B, Sellitto C, Cobellis G, Migliaccio M, Fasano S, Pierantoni R, Chianese R. Expression patterns of circular RNAs in high quality and poor quality human spermatozoa. *Front Endocrinol (Lausanne)* 2019;**10**:435.

Chirat F, Arkhis A, Martinage A, Jaquinod M, Chevaillier P, Sautière P. Phosphorylation of human sperm protamines HP1 and HP2: identification of phosphorylation sites. *Biochim Biophys Acta* 1993;**1203**:109–114.

Choi YH, Lim JK, Jeong MW, Kim KT. HnRNP A1 phosphorylated by VRK1 stimulates telomerase and its binding to telomeric DNA sequence. *Nucleic Acids Res* 2012;**40**:8499–8518.

Codina M, Estanyol JM, Fidalgo MJ, Ballescà JL, Oliva R. Advances in sperm proteomics: Best-practise methodology and clinical potential. *Expert Rev Proteomics* 2015;**12**:255–277.

Cohen DJ, Maldera JA, Vasen G, Ernesto JI, Munoz MW, Battistone MA, Cuasnicu PS. Epididymal Protein CRISP1 Plays Different Roles During the Fertilization Process. *J Androl* 2011;**32**:672–678.

Connolly MP, Hoorens S, Chambers GM. The costs and consequences of assisted reproductive technology: An economic perspective. *Hum Reprod Update* 2010;**16**:603–613.

Cooper TG. Cytoplasmic droplets: the good, the bad or just confusing? *Hum Reprod* 2005;**20**:9–11.

Cooper TG, Yeung CH. Computer-aided evaluation of assessment of “grade a” spermatozoa by experienced technicians. *Fertil Steril* 2006;**85**:220–224.

Cornwall GA. New insights into epididymal biology and function. *Hum Reprod Update* 2009;**15**:213–227.

Corzett M, Mazrimas J, Balhorn R. Protamine 1: Protamine 2 stoichiometry in the sperm of eutherian mammals. *Mol Reprod Dev* 2002;**61**:519–527.

da Ros VG, Muñoz MW, Battistone MA, Brukman NG, Carvajal G, Curci L, Gómez-Elías MD, Cohen DJ, Cuasnicu PS. From the epididymis to the egg: Participation of CRISP proteins in mammalian fertilization. *Asian J Androl* 2015;**17**:, p. 711–715.

Dacheux JL, Dacheux F. New insights into epididymal function in relation to sperm maturation. *Reproduction* 2014;**147**:147–174.

Dacheux JL, Dacheux F, Druart X. Epididymal protein markers and fertility. *Anim Reprod Sci* 2016;**169**:76–87.

Dang Y, Yan L, Hu B, Fan X, Ren Y, Li R, Lian Y, Yan J, Li Q, Zhang Y, *et al.* Tracing the expression of circular RNAs in human pre-implantation embryos. *Genome*

Biol 2016;**17**:130.

de Geyter C, Calhaz-Jorge C, Kupka MS, Wyns C, Mocanu E, Motrenko T, Scaravelli G, Smeenk J, Vidakovic S, Goossens V, *et al.* ART in Europe, 2014: Results generated from European registries by ESHRE. *Hum Reprod* 2018;**33**:1586–1601.

de Jonge C. Biological basis for human capacitation—revisited. *Hum Reprod Update* 2017;**23**:289–299.

de Kretser DM. Male infertility. *Lancet* 1997;**349**:787–790.

de Kretser DM, Loveland KL, Meinhardt A, Simorangkir D, Wreford N. Spermatogenesis. *Hum Reprod* 1998;**13**:, p. 1–8.

de Mateo S, Castillo J, Estanyol JM, Ballescà JL, Oliva R. Proteomic characterization of the human sperm nucleus. *Proteomics* 2011a;**11**:2714–2726.

de Mateo S, Estanyol JM, Oliva R. Methods for the analysis of the sperm proteome. *Methods Mol Biol* 2013;**927**:411–422.

de Mateo S, Gázquez C, Guimerà M, Balasch J, Meistrich ML, Ballescà JL, Oliva R. Protamine 2 precursors (pre-P2), protamine 1 to protamine 2 ratio (P1/P2), and assisted reproduction outcome. *Fertil Steril* 2009;**91**:715–722.

de Mateo S, Martínez-Heredia J, Estanyol JM, Domínguez-Fandos D, Domíguez-Fandos D, Vidal-Taboada JM, Ballescà JL, Oliva R. Marked correlations in protein expression identified by proteomic analysis of human spermatozoa. *Proteomics* 2007;**7**:4264–4277.

de Mateo S, Ramos L, de Boer P, Meistrich M, Oliva R. Protamine 2 precursors and processing. *Protein Pept Lett* 2011b;**18**:778–785.

de Mateo S, Ramos L, Van Der Vlag J, De Boer P, Oliva R. Improvement in chromatin maturity of human spermatozoa selected through density gradient centrifugation. *Int J Androl* 2011c;**34**:256–267.

de Yebra L, Ballesca JL, Vanrell JA, Bassas L, Oliva R. Complete selective absence of protamine P2 in humans. *J Biol Chem* 1993;**268**:10553–10557.

de Yebra L, Ballescá JL, Vanrell JA, Corzett M, Balhorn R, Oliva R. Detection of P2 precursors in the sperm cells of infertile patients who have reduced protamine P2 levels. *Fertil Steril* 1998;**69**:755–759.

de Yebra L, Oliva R. Rapid analysis of mammalian sperm nuclear proteins. *Anal Biochem* 1993;**209**:201–203.

Depa-Martynow M, Kempisty B, Jagodziński PP, Pawelczyk L, Jedrzejczak P. Impact of protamine transcripts and their proteins on the quality and fertilization ability of sperm and the development of preimplantation embryos. *Reprod Biol* 2012;**12**:57–72.

Depa-Martynów M, Kempisty B, Lianeri M, Jagodziński PP, Jędrzejczak P. Association between fertilin β, protamines 1 and 2 and spermatid-specific linker histone H1-like protein mRNA levels, fertilization ability of human spermatozoa, and quality of preimplantation embryos. *Folia Histochem Cytobiol* 2007;**45**:, p. 79–85.

Devroey P, Van Steirteghem A. A review of ten years experience of ICSI. *Hum Reprod Update* 2004;**10**:19–28.

Dey S, Brothag C, Vijayaraghavan S. Signaling Enzymes Required for Sperm Maturation and Fertilization in Mammals. *Front Cell Dev Biol* 2019;**7**:.

Dias BG, Ressler KJ. Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat Neurosci* 2014;**17**:89–96.

Diedrich K, Fauser BCJM, Devroey P, Griesinger G. The role of the endometrium and embryo in human implantation. *Hum Reprod Update* 2007;**13**:365–377.

Donkin I, Barrès R. Sperm epigenetics and influence of environmental factors. *Mol Metab* 2018;**14**:1–11.

Donkin I, Versteyhe S, Ingerslev LR, Qian K, Mechta M, Nordkap L, Mortensen B, Appel EVR, Jørgensen N, Kristiansen VB, *et al.* Obesity and bariatric surgery drive epigenetic variation of spermatozoa in humans. *Cell Metab* 2016;**23**:369–378.

Duncan FE, Que EL, Zhang N, Feinberg EC, O'Halloran T V., Woodruff TK. The zinc spark is an inorganic signature of human egg activation. *Sci Rep* 2016;**6**:1–8.

Enuka Y, Lauriola M, Feldman ME, Sas-Chen A, Ulitsky I, Yarden Y. Circular RNAs are long-lived and display only minimal early alterations in response to a growth factor. *Nucleic Acids Res* 2016;**44**:1370–1383.

Estill MS, Hauser R, Krawetz SA. RNA element discovery from germ cell to blastocyst. *Nucleic Acids Res* 2019;**47**:2263–2275.

Evers JLH. Female subfertility. *Lancet* 2002;**360**:, p. 151–159.

Ferrer-Vaquer A, Barragan M, Freour T, Vernaeve V, Vassena R. PLCζ sequence, protein levels, and distribution in human sperm do not correlate with semen characteristics and fertilization rates after ICSI. *J Assist Reprod Genet* 2016;**33**:747–756.

Ficarro S, Chertihin O, Westbrook VA, White F, Jayes F, Kalab P, Marto JA, Shabanowitz J, Herr JC, Hunt DF, *et al.* Phosphoproteome analysis of capacitated human sperm: Evidence of tyrosine phosphorylation of a kinase-anchoring protein 3 and valosin-containing protein/p97 during capacitation. *J Biol Chem* 2003;**278**:11579–11589.

Fishman EL, Jo K, Nguyen QPH, Kong D, Royfman R, Cekic AR, Khanal S, Miller AL, Simerly C, Schatten G, *et al.* A novel atypical sperm centriole is functional during human fertilization. *Nat Commun* 2018;**9**:2210.

Flaherty SP, Payne D, Matthews CD. *Fertilization failures and abnormal fertilization after intracytoplasmic sperm injection*. *Hum Reprod* 1998;**13**:.

Frapsauce C, Pionneau C, Bouley J, De Larouzière V, Berthaut I, Ravel C, Antoine JM, Soubrier F, Mandelbaum J. Infertilité masculine chez les patients normospermiques: Analyse protéomique des spermes normaux non fécondants en fécondation in vitro classique. *Gynecol Obstet Fertil* 2009;**37**:796–802.

Frapsauce C, Pionneau C, Bouley J, Delarouziere V, Berthaut I, Ravel C, Antoine JM, Soubrier F, Mandelbaum J. Proteomic identification of target proteins in normal but nonfertilizing sperm. *Fertil Steril* 2014;**102**:372–380.

Fujihara Y, Murakami M, Inoue N, Satouh Y, Kaseda K, Ikawa M, Okabe M. Sperm equatorial segment protein 1, SPESP1, is required for fully fertile sperm in mouse. *J Cell Sci* 2010;**123**:1531–1536.

Gadea J, Parrington J, Kashir J, Coward K. The male reproductive tract and spermatogenesis. *Textb Clin Embryol* 2013;, p. 18–26.

Galan C, Krykbaeva M, Rando OJ. Early life lessons: The lasting effects of germline epigenetic information on organismal development. *Mol Metab* 2020;**38**:100924.

Ganguly A, Bukovsky A, Sharma RK, Bansal P, Bhandari B, Gupta SK. In humans, zona pellucida glycoprotein-1 binds to spermatozoa and induces acrosomal exocytosis. *Hum Reprod* 2010;**25**:1643–1656.

Gapp K, Jawaid A, Sarkies P, Bohacek J, Pelczar P, Prados J, Farinelli L, Miska E, Mansuy IM. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat Neurosci* 2014;**17**:667–669.

García-Herrero S, Garrido N, Martínez-Conejero JA, Remohí J, Pellicer A, Meseguer M. Ontological evaluation of transcriptional differences between sperm of infertile males and fertile donors using microarray analysis. *J Assist Reprod Genet* 2010a;**27**:111–120.

García-Herrero S, Garrido N, Martínez-Conejero JA, Remohí J, Pellicer A, Meseguer M. Differential transcriptomic profile in spermatozoa achieving pregnancy or not via ICSI. *Reprod Biomed Online* 2011;**22**:25–36.

García-Herrero S, Meseguer M, Martínez-Conejero JA, Remohí J, Pellicer A, Garrido N. The transcriptome of spermatozoa used in homologous intrauterine insemination varies considerably between samples that achieve pregnancy and those that do not. *Fertil Steril* 2010b;**94**:1360–1373.

García-Peiró A, Martínez-Heredia J, Oliver-Bonet M, Abad C, Amengual MJ, Navarro

J, Jones C, Coward K, Gosálvez J, Benet J. Protamine 1 to protamine 2 ratio correlates with dynamic aspects of DNA fragmentation in human sperm. *Fertil Steril* 2011;**95**:105–109.

Garrido N, Martínez-Conejero JA, Jauregui J, Horcajadas JA, Simón C, Remohí J, Meseguer M. Microarray analysis in sperm from fertile and infertile men without basic sperm analysis abnormalities reveals a significantly different transcriptome. *Fertil Steril* 2009;**91**:1307–1310.

Gatewood JM, Cook GR, Balhorn R, Bradbury EM, Schmid CW. Sequence-specific packaging of DNA in human sperm chromatin. *Science (80-)* 1987;**236**:962–964.

Gázquez C, Oriola J, de Mateo S, Vidal-Taboada JM, Ballescà JL, Oliva R. A common protamine 1 promoter polymorphism (-190 C→{A) correlates with abnormal sperm morphlogy and increased protamine P1/P2 ratio in infertile patients. *J Androl* 2008;**29**:540–548.

Gervasi MG, Visconti PE. Molecular changes and signaling events occurring in spermatozoa during epididymal maturation. *Andrology* 2017;**5**:204–218.

Gòdia M, Castelló A, Rocco M, Cabrera B, Rodríguez-Gil JE, Balasch S, Lewis C, Sánchez A, Clop A. Identification of circular RNAs in porcine sperm and evaluation of their relation to sperm motility. *Sci Rep* 2020;**10**:1–11.

Gòdia M, Swanson G, Krawetz SA. A history of why fathers' RNA matters†. *Biol Reprod* 2018;**99**:147–159.

Goldberg E, Zirkin BR. Spermatogenesis: Overview. *Encycl Reprod* 2018;, p. 13–18.

Goodrich RJ, Anton E, Krawetz SA. Isolating mRNA and small noncoding RNAs from human sperm. *Methods Mol Biol* 2013;**927**:385–396.

Goring D, Indriolo E. Gene expression: How plants avoid incest. *Nature* 2010;**466**:926–928.

Gou LT, Lim DH, Ma W, Aubol BE, Hao Y, Wang X, Zhao J, Liang Z, Shao C, Zhang X, *et al.* Initiation of Parental Genome Reprogramming in Fertilized Oocyte by Splicing Kinase SRPK1-Catalyzed Protamine Phosphorylation. *Cell* 2020;**180**:1212-1227.e14.

Goudakou M, Kalogeraki A, Matalliotakis I, Panagiotidis Y, Gullo G, Prapas Y. Cryptic sperm defects may be the cause for total fertilization failure in oocyte donor cycles. *Reprod Biomed Online* 2012;**24**:148–152.

Govin J, Caron C, Lestrat C, Rousseaux S, Khochbin S. The role of histones in chromatin remodelling during mammalian spermiogenesis. *Eur J Biochem* 2004;**271**:3459–3469.

Grandjean V, Fourré S, De Abreu DAF, Derieppe MA, Remy JJ, Rassoulzadegan M.

RNA-mediated paternal heredity of diet-induced obesity and metabolic disorders. *Sci Rep* 2015;**5**:18193.

Green GR, Balhorn R, Poccia DL, Hecht NB. Synthesis and processing of mammalian protamines and transition proteins. *Mol Reprod Dev* 1994;**37**:255–263.

Greene J, Baird AM, Brady L, Lim M, Gray SG, McDermott R, Finn SP. Circular RNAs: Biogenesis, function and role in human diseases. *Front Mol Biosci* 2017;**6**:38.

Gu B, Zhang J, Wu Y, Zhang X, Tan Z, Lin Y, Huang X, Chen L, Yao K, Zhang M. Proteomic analyses reveal common promiscuous patterns of cell surface proteins on human embryonic stem cells and sperms. *PLoS One* 2011;**6**:e19386.

Guo L, Chao S-B, Xiao L, Wang Z-B, Meng T-G, Li Y-Y, Han Z-M, Ouyang Y-C, Hou Y, Sun Q-Y, *et al.* Sperm-carried RNAs play critical roles in mouse embryonic development. *Oncotarget* 2017;**8**:67394–67405.

Guo Y, Jiang W, Yu W, Niu X, Liu F, Zhou T, Zhang H, Li Y, Zhu H, Zhou Z, *et al.* Proteomics analysis of asthenozoospermia and identification of glucose-6-phosphate isomerase as an important enzyme for sperm motility. *J Proteomics* 2019;**208**:103478.

Gusse M, Sautiere P, Belaiche D, Martinage A, Roux C, Dadoune JP, Chevaillier P. Purification and characterization of nuclear basic proteins of human sperm. *Biochim Biophys Acta* 1986;**884**:124–134.

Gyamera-Acheampong C, Tantibhedhyangkul J, Weerachatyanukul W, Tadros H, Xu H, Van De Loo JW, Pelletier RM, Tanphaichitr N, Mbikay M. Sperm from mice genetically deficient for the PCSK4 proteinase exhibit accelerated capacitation, precocious acrosome reaction, reduced binding to egg zona pellucida, and impaired fertilizing ability. *Biol Reprod* 2006;**74**:666–673.

Hachem A, Godwin J, Ruas M, Lee HC, Buitrago MF, Ardestani G, Bassett A, Fox S, Navarrete F, De Sutter P, *et al.* Plcζ is the physiological trigger of the Ca2+ oscillations that induce embryogenesis in mammals but conception can occur in its absence. *Dev* 2017;**144**:2914–2924.

Hamada A, Esteves SC, Nizza M, Agarwal A. Unexplained male infertility: Diagnosis and management. *Int Braz J Urol* 2012;**38**:576–594.

Hammadeh ME, Hamad MF, Montenarh M, Fischer-Hammadeh C. Protamine contents and P1/P2 ratio in human spermatozoa from smokers and non-smokers. *Hum Reprod* 2010;**25**:2708–2720.

Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* 2009;**460**:473–478.

Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J.

Natural RNA circles function as efficient microRNA sponges. *Nature* 2013;**495**:384–388.

Hansen TB, Wiklund ED, Bramsen JB, Villadsen SB, Statham AL, Clark SJ, Kjems J. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J* 2011;**30**:4414–4422.

Hashemitabar M, Sabbagh S, Orazizadeh M, Ghadiri A, Bahmanzadeh M. A proteomic analysis on human sperm tail: Comparison between normozoospermia and asthenozoospermia. *J Assist Reprod Genet* 2015;**32**:853–863.

Hata T, Nakayama M. Targeted disruption of the murine large nuclear KIAA1440/Ints1 protein causes growth arrest in early blastocyst stage embryos and eventual apoptotic cell death. *Biochim Biophys Acta - Mol Cell Res* 2007;**1773**:1039–1051.

Heidary Z, Zaki-Dizaji M, Saliminejad K, Khorram Khorshid HR. MicroRNA profiling in spermatozoa of men with unexplained asthenozoospermia. *Andrologia* 2019;**51**:e13284.

Hernández-Silva G, Fabián López-Araiza JE, López-Torres AS, Larrea F, Torres-Flores V, Chirinos M. Proteomic characterization of human sperm plasma membrane-associated proteins and their role in capacitation. *Andrology* 2020;**8**:171–180.

Hetherington L, Schneider EK, DeKretser D, Muller CH, Hondermarck H, Velkov T, Baker MA. Deficiency in outer dense fiber 1 is a marker and potential driver of idiopathic male infertility. *Mol Cell Proteomics* 2016;**15**:3685–3693.

Heytens E, Parrington J, Coward K, Young C, Lambrecht S, Yoon SY, Fissore RA, Hamer R, Deane CM, Ruas M, *et al.* Reduced amounts and abnormal forms of phospholipase C zeta (PLCζ) in spermatozoa from infertile men. *Hum Reprod* 2009;**24**:2417–2428.

Ho K, Wolff CA, Suarez SS. CatSper-null mutant spermatozoa are unable to ascend beyond the oviductal reservoir. *Reprod Fertil Dev* 2009;**21**:345–350.

Holdcraft RW, Braun RE. Hormonal regulation of spermatogenesis. *Int J Androl* 2004;**27**:335–342.

Holstein AF, Schulze W, Davidoff M. Understanding spermatogenesis is a prerequisite for treatment. *Reprod Biol Endocrinol* 2003;**1**:.

Inoue D, Wittbrodt J, Gruss OJ. Loss and Rebirth of the Animal Microtubule Organizing Center: How Maternal Expression of Centrosomal Proteins Cooperates with the Sperm Centriole in Zygotic Centrosome Reformation. *BioEssays* 2018;**40**:e1700135.

Inoue N, Ikawa M, Isotani A, Okabe M. The immunoglobulin superfamily protein Izumo is required for sperm to fuse with eggs. *Nature* 2005;**434**:234–238.

Itoh K, Kondoh G, Miyachi H, Sugai M, Kaneko Y, Kitano S, Watanabe H, Maeda R, Imura A, Liu Y, *et al.* Dephosphorylation of protamine 2 at serine 56 is crucial for murine sperm maturation in vivo. *Sci Signal* 2019;**12**:eaa07232.

James ER, Carrell DT, Aston KI, Jenkins TG, Yeste M, Salas-Huetos A. The role of the epididymis and the contribution of epididymosomes to mammalian reproduction. *Int J Mol Sci* 2020;**21**:1–17.

Jeck WR, Sharpless NE. Detecting and characterizing circular RNAs. *Nat Biotechnol* 2014;**32**:453–461.

Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 2013;**19**:141–157.

Jodar M. Sperm and seminal plasma RNAs: What roles do they play beyond fertilization? *Reproduction* 2019;**158**:R113–R123.

Jodar M, Barrachina F, Oliva R. The Use of Sperm Proteomics in the Assisted Reproduction Laboratory. *A Pract Guid to Sperm Anal* 2018;233–244.

Jodar M, Kalko S, Castillo J, Ballescà JL, Oliva R. Differential RNAs in the sperm cells of asthenozoospermic patients. *Hum Reprod* 2012;**27**:1431–1438.

Jodar M, Oliva R. Protamine alterations in human spermatozoa. *Adv Exp Med Biol* 2014;**791**:83–102.

Jodar M, Selvaraju S, Sendler E, Diamond MP, Krawetz SA. The presence, role and clinical use of spermatozoal RNAs. *Hum Reprod Update* 2013;**19**:604–624.

Jodar M, Sendler E, Krawetz SA. The protein and transcript profiles of human semen. *Cell Tissue Res* 2016a;**363**:85–96.

Jodar M, Sendler E, Moskovtsev SI, Librach CL, Goodrich R, Swanson S, Hauser R, Diamond MP, Krawetz SA. Absence of sperm RNA elements correlates with idiopathic male infertility. *Sci Transl Med* 2015;**7**:295re6.

Jodar M, Sendler E, Moskovtsev SI, Librach CL, Goodrich R, Swanson S, Hauser R, Diamond MP, Krawetz SA. Response to Comment on "Absence of sperm RNA elements correlates with idiopathic male infertility." *Sci Transl Med* 2016b;**8**:.

Jodar M, Soler-Ventura A, Oliva R. Semen proteomics and male infertility. *J Proteomics* 2017;**6**:125–134.

Johnson GD, Mackie P, Jodar M, Moskovtsev S, Krawetz SA. Chromatin and extracellular vesicle associated sperm RNAs. *Nucleic Acids Res* 2015;**43**:6847–6859.

Johnson GD, Sendler E, Lalancette C, Hauser R, Diamond MP, Krawetz SA. Cleavage of rRNA ensures translational cessation in sperm at fertilization. *Mol Hum Reprod* 2011;**17**:721–726.

Johnston DS, Wooters J, Kopf GS, Qiu Y, Roberts KP. Analysis of the human sperm proteome. *Ann N Y Acad Sci* 2005;**1061**:, p. 190–202.

Jones EL, Zalensky AO, Zalenskaya IA. Protamine Withdrawal from Human Sperm Nuclei Following Heterologous ICSI into Hamster Oocytes. *Protein Pept Lett* 2012;**18**:811–816.

Jumeau F, Com E, Lane L, Duek P, Lagarrigue M, Lavigne R, Guillot L, Rondel K, Gateau A, Melaine N, *et al.* Human spermatozoa as a model for detecting missing proteins in the context of the chromosome-centric human proteome project. *J Proteome Res* 2015;**14**:3606–3620.

Kaati G, Bygren LO, Edvinsson S. Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. *Eur J Hum Genet* 2002;**10**:682–688.

Kashir J, Konstantinidis M, Jones C, Heindryckx B, De Sutter P, Parrington J, Wells D, Coward K. Characterization of two heterozygous mutations of the oocyte activation factor phospholipase C zeta (PLCζ) from an infertile man by use of minisequencing of individual sperm and expression in somatic cells. *Fertil Steril* 2012;**98**:423–431.

Kelleher NL, Thomas PM, Ntai I, Compton PD, Leduc RD. Deep and quantitative top-down proteomics in clinical and translational research. *Expert Rev Proteomics* 2014;**11**:649–651.

Khara KK. Human protamines and male infertility. *J Assist Reprod Genet* 1997;**14**:282–290.

Kichine E, Di Falco M, Hales BF, Robaire B, Chan P. Analysis of the sperm head protein profiles in fertile men: consistency across time in the levels of expression of heat shock proteins and peroxiredoxins. *PLoS One* 2013;**8**:.

Kierszenbaum AL. Tyrosine protein kinases and spermatogenesis: Truncation matters. *Mol Reprod Dev* 2006;**73**:399–403.

Kim YH, Haidl G, Schaefer M, Egner U, Mandal A, Herr JC. Compartmentalization of a unique ADP/ATP carrier protein SFEC (Sperm Flagellar Energy Carrier, AAC4) with glycolytic enzymes in the fibrous sheath of the human sperm flagellar principal piece. *Dev Biol* 2007;**302**:463–476.

Kimmins S, Sassone-Corsi P. Chromatin remodelling and epigenetic features of germ cells. *Nature* 2005;**434**:583–589.

Kleaveland B, Shi CY, Stefano J, Bartel DP. A Network of Noncoding Regulatory

RNAs Acts in the Mammalian Brain. *Cell* 2018;**174**:350-362.e17.

Kramer IjM. Protein Phosphatases. *Signal Transduct* 2016;, p. 935–995.

Kramer MC, Liang D, Tatomer DC, Gold B, March ZM, Cherry S, Wilusz JE. Combinatorial control of *Drosophila* circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes Dev* 2015;**29**:2168–2182.

Krausz C. Male infertility: Pathogenesis and clinical diagnosis. *Best Pract Res Clin Endocrinol Metab* 2011;**25**:271–285.

Krausz C, Riera-Escamilla A. Genetics of male infertility. *Nat Rev Urol* 2018;**15**:369–384.

Krawetz S a. Paternal contribution: new insights and future challenges. *Nat Rev Genet* 2005;**6**:633–642.

Krawetz SA, Kruger A, Lalancette C, Tagett R, Anton E, Draghici S, Diamond MP. A survey of small RNAs in human sperm. *Hum Reprod* 2011;**26**:3401–3412.

Krejčí J, Stixová L, Pagáčová E, Legartová S, Kozubek S, Lochmanová G, Zdráhal Z, Sehnalová P, Dabravolski S, Hejátko J, *et al.* Post-Translational Modifications of Histones in Human Sperm. *J Cell Biochem* 2015;**116**:2195–2209.

Kristensen LS, Andersen MS, Stagsted LVW, Ebbesen KK, Hansen TB, Kjems J. The biogenesis, biology and characterization of circular RNAs. *Nat Rev Genet* 2019;**20**:675–691.

Kruger TF, Acosta AA, Simmons KF, Swanson RJ, Matta JF, Veeck LL, Morshedi M, Brugo S. New method of evaluating sperm morphology with predictive value for human in vitro fertilization. *Urology* 1987;**30**:248–251.

Kvist U, Björndahl L. Zinc preserves an inherent capacity for human sperm chromatin decondensation. *Acta Physiol Scand* 1985;**124**:195–200.

Lefièvre L, Chen Y, Conner SJ, Scott JL, Publicover SJ, Ford WCL, Barratt CLR. Human spermatozoa contain multiple targets for protein S-nitrosylation: An alternative mechanism of the modulation of sperm function by nitric oxide? *Proteomics* 2007;**7**:3066–3084.

Légaré C, Droit A, Fournier F, Bourassa S, Force A, Cloutier F, Tremblay R, Sullivan R. Investigation of male infertility using quantitative comparative proteomics. *J Proteome Res* 2014;**13**:5403–5414.

Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M, *et al.* Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Mol Cell* 2017;**66**:22-37.e9.

Lescoat D, Colleu D, Boujard D, Le Lannou D. Electrophoretic characteristics of nuclear proteins from human spermatozoa. *Arch Androl* 1988;**20**:35–40.

Li LW, Fan LQ, Zhu WB, Nie HC, Sun BL, Luo KL, Liao TT, Tang L, Lu GX. Establishment of a high-resolution 2-D reference map of human spermatozoal proteins from 12 fertile sperm-bank donors. *Asian J Androl* 2007;**9**:321–329.

Liao TT, Xiang Z, Zhu WB, Fan LQ. Proteome analysis of round-headed and normal spermatozoa by 2-D fluorescence difference gel electrophoresis and mass spectrometry. *Asian J Androl* 2009;**11**:683–693.

Lin JB, Troyer D. Testicular Anatomy and Physiology. *Pathobiol Hum Dis A Dyn Encycl Dis Mech* 2014;, p. 2464–2475.

Liu M. The biology and dynamics of mammalian cortical granules. *Reprod Biol Endocrinol* 2011;**17**:149.

Liu W-M, Pang RTK, Chiu PCN, Wong BPC, Lao K, Lee K-F, Yeung WSB. Sperm-borne microRNA-34c is required for the first cleavage division in mouse. *Proc Natl Acad Sci* 2012;**109**:490–494.

Liu X, Liu G, Liu J, Zhu P, Wang J, Wang Y, Wang W, Li N, Wang X, Zhang C, *et al.* ITRAQ-based analysis of sperm proteome from normozoospermic men achieving the rescue-ICSI pregnancy after the IVF failure. *Clin Proteomics* 2018;**15**:27.

Liu Y, Guo Y, Song N, Fan Y, Li K, Teng X, Guo Q, Ding Z. Proteomic pattern changes associated with obesity-induced asthenozoospermia. *Andrology* 2015;**3**:247–259.

Lou C, Goodier JL, Qiang R. A potential new mechanism for pregnancy loss: Considering the role of LINE-1 retrotransposons in early spontaneous miscarriage. *Reprod Biol Endocrinol* 2020;**18**:6.

Luense LJ, Wang X, Schon SB, Weller AH, Lin Shiao E, Bryant JM, Bartolomei MS, Coutifaris C, Garcia BA, Berger SL. Comprehensive analysis of histone post-translational modifications in mouse and human male germ cells. *Epigenetics and Chromatin* 2016;**21**:24.

Luke B, Brown MB, Wantman E, Lederman A, Gibbons W, Schattman GL, Lobo RA, Leach RE, Stern JE. Cumulative Birth Rates with Linked Assisted Reproductive Technology Cycles. *N Engl J Med* 2012;**366**:2483–2491.

Maher ER. Imprinting and assisted reproductive technology. *Hum Mol Genet* 2005;**15**:R133-8.

Maldera JA, Muñoz MW, Chirinos M, Busso D, Raffo FGE, Battistone MA, Blaquier JA, Larrea F, Cuasnicu PS. Human fertilization: Epididymal hCRISP1 mediates sperm-zona pellucida binding through its interaction with ZP3. *Mol Hum Reprod* 2014;**20**:341–349.

Manandhar G, Schatten H, Sutovsky P. Centrosome reduction during gametogenesis and its significance. *Biol Reprod* 2005;**72**:2–13.

Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA, Ferrari R. Genome, transcriptome and proteome: The rise of omics data and their integration in biomedical sciences. *Brief Bioinform* 2018;**19**:286–302.

Mao S, Goodrich RJ, Hauser R, Schrader SM, Chen Z, Krawetz SA. Evaluation of the effectiveness of semen storage and sperm purification methods for spermatozoa transcript profiling. *Syst Biol Reprod Med* 2013;**59**:287–295.

Mao S, Sendler E, Goodrich RJ, Hauser R, Krawetz SA. A comparison of sperm RNA-seq methods. *Syst Biol Reprod Med* 2014;**60**:308–315.

Martianov I, Brancorsini S, Catena R, Gansmuller A, Kotaja N, Parvinen M, Sassone-Corsi P, Davidson I. Polar nuclear localization of H1T2, a histone H1 variant, required for spermatid elongation and DNA condensation during spermiogenesis. *Proc Natl Acad Sci U S A* 2005;**102**:2808–2813.

Martínez-Heredia J, de Mateo S, Vidal-Taboada JM, Ballescà JL, Oliva R. Identification of proteomic differences in asthenozoospermic sperm samples. *Hum Reprod* 2008;**23**:783–791.

Martínez-Heredia J, Estanyol JM, Ballescà JL, Oliva R. Proteomic identification of human sperm proteins. *Proteomics* 2006;**6**:4356–4369.

Maze I, Noh KM, Soshnev AA, Allis CD. Every amino acid matters: Essential contributions of histone variants to mammalian development and disease. *Nat Rev Genet* 2014;**15**:259–271.

McLachlan RI, Yazdani A, Kovacs G, Howlett D. Management of the infertile couple. *Aust Fam Physician* 2005;**34**:111–117.

McReynolds S, Dzieciatkowska M, Stevens J, Hansen KC, Schoolcraft WB, Katz-Jaffe MG. Toward the identification of a subset of unexplained infertility: A sperm proteomic approach. *Fertil Steril* 2014;**102**:692–699.

Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, *et al.* Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 2013;**495**:333–338.

Mengual L, Ballescà JL, Ascaso C, Oliva R. Marked Differences in Protamine Content and P1 / P2 Ratios. *J Androl* 2003;**24**:438–447.

Mezquita C. Chromatin composition, structure and function in spermatogenesis. *Revis Biol Celular* 1985;**5**:.

Miller D. Confrontation, consolidation, and recognition: The oocyte's perspective on the incoming sperm. *Cold Spring Harb Perspect Med* 2015;**5**:1–18.

Miller D, Iles D. Has the renewed interest in sperm RNA led to fresh insights? A critical review and hypothesis. In Carrell DT, editor. *Paternal Influ Hum Reprod*

Success 2013;, p. 38–49.

Misell LM, Holochwost D, Boban D, Santi N, Shefi S, Hellerstein MK, Turek PJ. A stable isotope-mass spectrometric method for measuring human spermatogenesis kinetics in vivo. *J Urol* 2006;**175**:242–246.

Molina LCP, Luque GM, Balestrini PA, Marín-Briggiler CI, Romarowski A, Buffone MG. Molecular basis of human sperm capacitation. *Front Cell Dev Biol* 2018;**6**:72.

Montjean D, De La Grange P, Gentien D, Rapinat A, Belloc S, Cohen-Bacrie P, Menezo Y, Benkhalifa M. Sperm transcriptome profiling in oligozoospermia. *J Assist Reprod Genet* 2012;**29**:3–10.

Mortimer D, Menkveld R. Sperm morphology assessment - Historical perspectives and current opinions. *J Androl* 2001;**22**:192–205.

Moscatelli N, Lunetti P, Braccia C, Armirotti A, Pisanello F, De Vittorio M, Zara V, Ferramosca A. Comparative proteomic analysis of proteins involved in bioenergetics pathways associated with human sperm motility. *Int J Mol Sci* 2019;**20**:3000.

Naaby-Hansen S, Diekman A, Shetty J, Flickinger CJ, Westbrook A, Herr JC. Identification of calcium-binding proteins associated with the human sperm plasma membrane. *Reprod Biol Endocrinol* 2010;**8**:.

Naaby-Hansen S, Herr JC. Heat shock proteins on the human sperm surface. *J Reprod Immunol* 2010;**84**:32–40.

Nanassy L, Liu L, Griffin J, Carrell DT. The clinical utility of the protamine 1/protamine 2 ratio in sperm. *Protein Pept Lett* 2011;**18**:772–777.

Nasr-Esfahani MH, Razavi S, Mozdarani H, Mardani M, Azvagi H. Relationship between protamine deficiency with fertilization rate and incidence of sperm premature chromosomal condensation post-ICSI. *Andrologia* 2004a;**36**:95–100.

Nasr-Esfahani MH, Salehi M, Razavi S, Mardani M, Bahramian H, Steger K, Oreizi F. Effect of protamine-2 deficiency on ICSI outcome. *Reprod Biomed Online* 2004b;**9**:652–658.

Nelson JE, Krawetz SA. Mapping the clonally unstable recombinogenic PRM1→PRM2→TNP2 region of human 16p 13.2. *Mitochondrial DNA* 1995;**5**:163–168.

Ni K, Spiess A-N, Schuppe H-C, Steger K. The impact of sperm protamine deficiency and sperm DNA damage on human male fertility: a systematic review and meta-analysis. *Andrology* 2016;**4**:789–799.

Niakan KK, Han J, Pedersen RA, Simon C, Pera RAR. Human pre-implantation embryo development. *Development* 2012;**139**:829–841.

Niederberger C. Male Infertility. *J Urol* 2009;**182**:1510–1513.

Nixon B, De Iuliis GN, Hart HM, Zhou W, Mathe A, Bernstein IR, Anderson AL, Stanger SJ, Skerrett-Byrne DA, Fairuz Jamaluddin MB, *et al.* Proteomic profiling of mouse epididymosomes reveals their contributions to post-testicular sperm maturation. *Mol Cell Proteomics* 2019;**18**:S91–S108.

Nixon B, Mitchell LA, Anderson AL, Mclaughlin EA, O'bryan MK, Aitken RJ. Proteomic and functional analysis of human sperm detergent resistant membranes. *J Cell Physiol* 2011;**226**:2651–2665.

Noda T, Lu Y, Fujihara Y, Oura S, Koyano T, Kobayashi S, Matzuk MM, Ikawa M. Sperm proteins SOF1, TMEM95, and SPACA6 are required for sperm-oocyte fusion in mice. *Proc Natl Acad Sci U S A* 2020;**117**:11493–11502.

Nowicka-Bauer K, Lepczynski A, Ozgo M, Kamieniczna M, Fraczek M, Stanski L, Olszewska M, Malcher A, Skrzypczak W, Kurpisz MK. Sperm mitochondrial dysfunction and oxidative stress as possible reasons for isolated asthenozoospermia. *J Physiol Pharmacol* 2018;**69**:.

Nozawa K, Satouh Y, Fujimoto T, Oji A, Ikawa M. Sperm-borne phospholipase C zeta-1 ensures monospermic fertilization in mice. *Sci Rep* 2018;**8**:1315.

Ntostis P, Carter D, Iles D, Huntriss J, Tzetis M, Miller D. Potential sperm contributions to the murine zygote predicted by in silico analysis. *Reproduction* 2017;**154**:777–788.

Oh JS, Han C, Cho C. ADAM7 is associated with epididymosomes and integrated into sperm plasma membrane. *Mol Cells* 2009;**28**:441–446.

Ohto U, Ishida H, Krayukhina E, Uchiyama S, Inoue N, Shimizu T. Structure of IZUMO1-JUNO reveals sperm-oocyte recognition during mammalian fertilization. *Nature* 2016;**534**:566–569.

Okada Y, Yamagata K, Hong K, Wakayama T, Zhang Y. A role for the elongator complex in zygotic paternal genome demethylation. *Nature* 2010;**463**:554–558.

Oliva R. Sequence, evolution and transcriptional regulation of mammalian P1 type protamines. In Jamieson B, Ausió J, Justine J, editors. *Adv spermatozoal phylogeny Taxon* 1995;, p. 166, 537–548.

Oliva R. Protamines and male infertility. *Hum Reprod Update* 2006;**12**:417–435.

Oliva R, Ballescà JL. Altered histone retention and epigenetic modifications in the sperm of infertile men. *Asian J Androl* 2012;**14**:239–240.

Oliva R, Bazett-Jones D, Mezquita C, Dixon GH. Factors affecting nucleosome disassembly by protamines in vitro. Histone hyperacetylation and chromatin structure, time dependence, and the size of the sperm nuclear proteins. *J Biol*

Chem 1987;**262**:17016–17025.

Oliva R, Castillo J. Proteomics and the genetics of sperm chromatin condensation. *Asian J Androl* 2011;**13**:24–30.

Oliva R, de Mateo S, Estanyol JM. Sperm cell proteomics. *Proteomics* 2009;**9**:1004–1017.

Oliva R, Dixon GH. Vertebrate Protamine Genes and the Histone-to-Protamine Replacement Reaction. *Prog Nucleic Acid Res Mol Biol* 1991;**40**:25–94.

Oliva R, Martínez-Heredia J, Estanyol JM. Proteomics in the study of the sperm cell composition, differentiation and function. *Syst Biol Reprod Med* 2008;**54**:23–36.

Ostermeier GC, Dix DJ, Miller D, Khatri P, Krawetz SA. Spermatozoal RNA profiles of normal fertile men. *Lancet* 2002;**360**:772–777.

Ostermeier GC, Goodrich RJ, Moldenhauer JS, Diamond MP, Krawetz SA. A suite of novel human spermatozoal RNAs. *J Androl* 2005;**26**:70–74.

Ostermeier GC, Miller D, Huntriss JD, Diamond MP, Krawetz SA. Delivering spermatozoan RNA to the oocyte. *Nature* 2004;**429**:154–154.

Ozsolak F, Milos PM. RNA sequencing: Advances, challenges and opportunities. *Nat Rev Genet* 2011;**12**:87–98.

Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E, *et al.* Translation of CircRNAs. *Mol Cell* 2017;**66**:9-21.e7.

Pan MM, Hockenberry MS, Kirby EW, Lipshultz LI. Male Infertility Diagnosis and Treatment in the Era of In Vitro Fertilization and Intracytoplasmic Sperm Injection. *Med Clin North Am* 2018;**102**:337–347.

Panner Selvam MK, Baskaran S, Agarwal A, Henkel R. Protein profiling in unlocking the basis of varicocele-associated infertility. *Andrologia* 2020;e13645.

Pantano L, Jodar M, Bak M, Ballesca JL, Tommerup N, Oliva R, Vavouri T. The small RNA content of human sperm reveals pseudogene-derived piRNAs complementary to protein-coding genes. *Rna* 2015;**21**:1085–1095.

Papoutsopoulou S, Nikolakaki E, Chalepakis G, Kruft V, Chevaillier P, Giannakouros T. SR protein-specific kinase 1 is highly expressed in testis and phosphorylates protamine 1. *Nucleic Acids Res* 1999;**27**:2972–2980.

Pappireddi N, Martin L, Wühr M. A Review on Quantitative Multiplexed Proteomics. *ChemBioChem* 2019;**20**:1210–1224.

Park OH, Ha H, Lee Y, Boo SH, Kwon DH, Song HK, Kim YK. Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex. *Mol Cell* 2019;**74**:494-507.e8.

Parte PP, Rao P, Redij S, Lobo V, D'Souza SJ, Gajbhiye R, Kulkarni V. Sperm phosphoproteome profiling by ultra performance liquid chromatography followed by data independent analysis (LC-MSE) reveals altered proteomic signatures in asthenozoospermia. *J Proteomics* 2012;**75**:5861–5871.

Patil S, Totey S. Developmental failure of hybrid embryos generated by in vitro fertilization of water buffalo (Bubalus bubalis) oocyte with bovine spermatozoa. *Mol Reprod Dev* 2003;**64**:360–368.

Patop IL, Wüst S, Kadener S. Past, present, and future of circRNAs. *EMBO J* 2019;**38**:e100836.

Patrat C, Serres C, Jouannet P. The acrosome reaction in human spermatozoa. *Biol Cell* 2000;**92**:255–266.

Peng H, Shi J, Zhang Y, Zhang H, Liao S, Li W, Lei L, Han C, Ning L, Cao Y, *et al.* A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. *Cell Res* 2012;**22**:1609–1612.

Perez MF, Lehner B. Intergenerational and transgenerational epigenetic inheritance in animals. *Nat Cell Biol* 2019;**21**:143–151.

Pini T, Parks J, Russ J, Dzieciatkowska M, Hansen KC, Schoolcraft WB, Katz-Jaffe M. Obesity significantly alters the human sperm proteome, with potential implications for fertility. *J Assist Reprod Genet* 2020;**37**:777–787.

Pirhonen A, Linnala-Kankkunen A, Mäenpää PH. Identification of Phosphoseryl Residues in Protamines from Mature Mammalian Spermatozoa1. *Biol Reprod* 1994a;**50**:981–986.

Pirhonen A, Linnala-Kankkunen A, Mäenpää PH. Identification of Phosphoseryl Residues in Protamines from Mature Mammalian Spermatozoa1. *Biol Reprod* 1994b;**50**:981–986.

Pirhonen A, Linnala-Kankkunen A, Mënpää PH. P2 protamines are phosphorylated in vitro by protein kinase C, whereas P1 protamines prefer cAMP-dependent protein kinase. A comparative study of five mammalian species. *Eur J Biochem* 1994c;**223**:165–169.

Pixton KL, Deeks ED, Flesch FM, Moseley FLC, Björndahl L, Ashton PR, Barratt CLR, Brewis IA. Sperm proteome mapping of a patient who experiencd failed fertilization at IVF reveals altered expression of at least 20 proteins compared with fertile donors: Case report. *Hum Reprod* 2004;**19**:1438–1447.

Plant TM. The hypothalamo-pituitary-gonadal axis. *J Endocrinol* 2015;**226**:T41–T54.

Platts AE, Dix DJ, Chemes HE, Thompson KE, Goodrich R, Rockett JC, Rawe VY, Quintana S, Diamond MP, Strader LF, *et al.* Success and failure in human spermatogenesis as revealed by teratozoospermic RNAs. *Hum Mol Genet*

2007;**16**:763–773.

Pruslin FH, Imesch E, Winston R, Rodman TC. Phosphorylation state of protamines 1 and 2 in human spermatids and spermatozoa. *Gamete Res* 1987;**18**:179–190.

Queralt R, Adroer R, Oliva R, Winkfein RJ, Retief JD, Dixon GH. Evolution of protamine P1 genes in mammals. *J Mol Evol* 1995;**40**:601–607.

Ragusa M, Barbagallo D, Chioccarelli T, Manfrevola F, Cobellis G, Di Pietro C, Brex D, Battaglia R, Fasano S, Ferraro B, *et al.* CircNAPEPLD is expressed in human and murine spermatozoa and physically interacts with oocyte miRNAs. *RNA Biol* 2019;**16**:1237–1248.

Ramadan WM, Kashir J, Jones C, Coward K. Oocyte activation and phospholipase C zeta (PLCζ): Diagnostic and therapeutic implications for assisted reproductive technology. *Cell Commun Signal* 2012;**10**:12.

Rando OJ. Daddy issues: Paternal effects on phenotype. *Cell* 2012;**151**:702–708.

Rassoulzadegan M, Grandjean V, Gounon P, Vincent S, Gillot I, Cuzin F. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* 2006;**441**:469–474.

Raya Á, Izpisúa Belmonte JC. Left-right asymmetry in the vertebrate embryo: From early information to higher-level integration. *Nat Rev Genet* 2006;**7**:283–293.

Redgrove KA, Anderson AL, Dun MD, McLaughlin EA, O'Bryan MK, Aitken RJ, Nixon B. Involvement of multimeric protein complexes in mediating the capacitation-dependent binding of human spermatozoa to homologous zonae pellucidae. *Dev Biol* 2011;**356**:460–474.

Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. *Science (80-)* 2001;**293**:1089–1093.

Reik W, Walter J. Genomic imprinting: Parental influence on the genome. *Nat Rev Genet* 2001;**2**:21–32.

Ribas-Maynou J, García-Peiró A, Martínez-Heredia J, Fernández-Encinas A, Abad C, Amengual MJ, Navarro J, Benet J. Nuclear degraded sperm subpopulation is affected by poor chromatin compaction and nuclease activity. *Andrologia* 2015;**47**:286–294.

Ritchie MD, Holzinger ER, Li R, Pendergrass SA, Kim D. Methods of integrating data to uncover genotype-phenotype interactions. *Nat Rev Genet* 2015;**16**:85–97.

Rodgers AB, Morgan CP, Leu NA, Bale TL. Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. *Proc Natl Acad Sci U S A* 2015;**112**:13699–13704.

Rogenhofer N, Dansranjavin T, Schorsch M, Spiess A, Wang H, Von Schönfeldt V,

Cappallo-Obermann H, Baukloh V, Yang H, Paradowska A, *et al.* The sperm protamine mRNA ratio as a clinical parameter to estimate the fertilizing potential of men taking part in an ART programme. *Hum Reprod* 2013;**28**:969–978.

Rogenhofer N, Ott J, Pilatz A, Wolf J, Thaler CJ, Windischbauer L, Schagdarsurengin U, Steger K, von Schönfeldt V. Unexplained recurrent miscarriages are associated with an aberrant sperm protamine mRNA content. *Hum Reprod* 2017;**32**:1574–1582.

Rousseaux S, Caron C, Govin J, Lestrat C, Faure AK, Khochbin S. Establishment of male-specific epigenetic information. *Gene* 2005;**345**:139–153.

Royo H, Stadler MB, Peters AHFM. Alternative Computational Analysis Shows No Evidence for Nucleosome Enrichment at Repetitive Sequences in Mammalian Spermatozoa. *Dev Cell* 2016;**37**:98–104.

Sala P, Ferrero S, Buffi D, Pastorino D, Bertoldi S, Vaccari L, Bentivoglio G, Venturini PL, De Biasio P. Congenital defects in assisted reproductive technology pregnancies. *Minerva Ginecol* 2011;**63**:227–22735.

Salas-Huetos A, Blanco J, Vidal F, Godo A, Grossmann M, Pons MC, F-Fernández S, Garrido N, Anton E. Spermatozoa from patients with seminal alterations exhibit a differential micro-ribonucleic acid profile. *Fertil Steril* 2015;**104**:591–601.

Salas-Huetos A, Blanco J, Vidal F, Grossmann M, Pons MC, Garrido N, Anton E. Spermatozoa from normozoospermic fertile and infertile individuals convey a distinct miRNA cargo. *Andrology* 2016;**4**:1028–1036.

Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. In Preiss T, editor. *PLoS One* 2012;**7**:e30733.

Santos F, Dean W. Epigenetic reprogramming during early development in mammals. *Reproduction* 2004;**127**:643–651.

Saraswat M, Joenväärä S, Jain T, Tomar AK, Sinha A, Singh S, Yadav S, Renkonen R. Human spermatozoa quantitative proteomic signature classifies normo-and asthenozoospermia. *Mol Cell Proteomics* 2017;**16**:57–72.

Satouh Y, Ikawa M. New Insights into the Molecular Events of Mammalian Fertilization. *Trends Biochem Sci* 2018;**43**:818–828.

Schagdarsurengin U, Steger K. Epigenetics in male reproduction: effect of paternal diet on sperm quality and offspring health. *Nat Rev Urol* 2016;**13**:584–595.

Schatten G. The centrosome and its mode of inheritance: The reduction of the centrosome during gametogenesis and its restoration during fertilization. *Dev Biol* 1994;**165**:299–335.

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T. The RIN: An RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 2006;**31**:3.

Schuster A, Tang C, Xie Y, Ortogero N, Yuan S, Yan W. SpermBase: A Database for Sperm-Borne RNA Contents. *Biol Reprod* 2016;**95**:99–99.

Secciani F, Bianchi L, Ermini L, Cianti R, Armini A, La Sala GB, Focarelli R, Bini L, Rosati F. Protein profile of capacitated versus ejaculated human sperm. *J Proteome Res* 2009;**8**:3377–3389.

Selvam MKP, Agarwal A, Pushparaj PN, Baskaran S, Bendou H. Sperm proteome analysis and identification of fertility-associated biomarkers in unexplained male infertility. *Genes (Basel)* 2019;**10**:522.

Sendler E, Johnson GD, Mao S, Goodrich RJ, Diamond MP, Hauser R, Krawetz SA. Stability, delivery and functions of human sperm RNAs at fertilization. *Nucleic Acids Res* 2013;**41**:4104–4117.

Shafei RA, Syrkasheva AG, Romanov AY, Makarova NP, Dolgushina N V., Semenova ML. Blastocyst hatching in humans. *Russ J Dev Biol* 2017;**48**:5–15.

Sharma R, Agarwal A. Spermatogenesis: An Overview. *Sperm Chromatin* 2011;, p. 19–44.

Sharma U, Conine CC, Shea JM, Boskovic A, Derr AG, Bing XY, Belleannee C, Kucukural A, Serra RW, Sun F, *et al.* Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* 2016;**351**:391–396.

Sharma U, Rando OJ. Father-son chats: Inheriting stress through sperm RNA. *Cell Metab* 2014;**19**:894–895.

Shen S, Wang J, Liang J, He D. Comparative proteomic study between human normal motility sperm and idiopathic asthenozoospermia. *World J Urol* 2013;**31**:1395–1401.

Shi S, Shi Q, Sun Y. The effect of sperm miR-34c on human embryonic development kinetics and clinical outcomes. *Life Sci* 2020;**256**:117895.

Siklenka K, Erkek S, Godmann M, Lambrot R, McGraw S, Lafleur C, Cohen T, Xia J, Suderman M, Hallett M, *et al.* Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. *Science* 2015;**350**:aab2006.

Simon L, Castillo J, Oliva R, Lewis SEM. Relationships between human sperm protamines, DNA damage and assisted reproduction outcomes. *Reprod Biomed Online* 2011;**23**:724–734.

Simon L, Liu L, Murphy K, Ge S, Hotaling J, Aston KI, Emery B, Carrell DT.

Comparative analysis of three sperm DNA damage assays and sperm nuclear protein content in couples undergoing assisted reproduction treatment. *Hum Reprod* 2014;**29**:904–917.

Singh AP, Rajender S. CatSper channel, sperm function and male fertility. *Reprod Biomed Online* 2015;**30**:28–38.

Sinha A, Singh V, Singh S, Yadav S. Proteomic analyses reveal lower expression of TEX40 and ATP6V0A2 proteins related to calcium ion entry and acrosomal acidification in asthenozoospermic males. *Life Sci* 2019;**218**:81–88.

Siva AB, Kameshwari DB, Singh V, Pavani K, Sundaram CS, Rangaraj N, Deenadayal M, Shivaji S. Proteomics-based study on asthenozoospermia: Differential expression of proteasome alpha complex. *Mol Hum Reprod* 2010;**16**:452–462.

Smith LM, Kelleher NL. Proteoforms as the next proteomics currency. *Science* 2018;**359**:1106–1107.

Smith LM, Kelleher NL, Consortium for Top Down Proteomics. Proteoform: a single term describing protein complexity. *Nat Methods* 2013;**10**:186–187.

Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung L-H, Bindereif A. Exon Circularization Requires Canonical Splice Signals. *Cell Rep* 2015;**10**:103–111.

Steger K, Wilhelm J, Konrad L, Stalf T, Greb R, Diemer T, Kliesch S, Bergmann M, Weidner W. Both protamine-1 to protamine-2 mRNA ratio and Bcl2 mRNA content in testicular spermatids and ejaculated spermatozoa discriminate between fertile and infertile men. *Hum Reprod* 2008;**23**:11–16.

Suarez SS, Wolfner MF. Seminal plasma plays important roles in fertility. *Sperm Cell Prod Matur Fertil Regen* 2017;, p. 88–108.

Sullivan R, Belleannée C. Role of the epididymis in sperm maturation. *Sperm Cell Prod Matur Fertil Regen* 2017;, p. 73–87.

Sullivan R, Mieusset R. The human epididymis: Its function in sperm maturation. *Hum Reprod Update* 2016;**22**:574–587.

Sullivan R, Saez F. Epididymosomes, prostasomes, and liposomes: Their roles in mammalian male reproductive physiology. *Reproduction* 2013;**146**:R21–R35.

Sun G, Jiang M, Zhou T, Guo Y, Cui Y, Guo X, Sha J. Insights into the lysine acetylproteome of human sperm. *J Proteomics* 2014;**109**:199–211.

Svejstrup JQ. Elongator complex: how many roles does it play? *Curr Opin Cell Biol* 2007;**19**:331–336.

Swanson GM, Estill M, Diamond MP, Legro RS, Coutifaris C, Barnhart KT, Huang H, Hansen KR, Trussell JC, Coward RM, *et al.* Human chromatin remodeler cofactor, RNA interactor, eraser and writer sperm RNAs responding to obesity.

Epigenetics 2020;**15**:32–46.

Szabo L, Salzman J. Detecting circular RNAs: bioinformatic and experimental challenges. *Nat Rev Genet* 2016;**17**:679–692.

Tabibzadeh S. Isolation, characterization, and function of EBAF/LEFTY B: Role in infertility. *Ann N Y Acad Sci* 2011;**1221**:98–102.

Tanaka H, Iguchi N, Isotani A, Kitamura K, Toyama Y, Matsuoka Y, Onishi M, Masai K, Maekawa M, Toshimori K, *et al.* HANP1/H1T2, a Novel Histone H1-Like Protein Involved in Nuclear Formation and Sperm Fertility. *Mol Cell Biol* 2005;**25**:7107–7119.

Tang C, Xie Y, Yu T, Liu N, Wang Z, Woolsey RJ, Tang Y, Zhang X, Qin W, Zhang Y, *et al.* m6A-dependent biogenesis of circular RNAs in male germ cells. *Cell Res* 2020;**30**:211–228.

Tardif S, Guyonnet B, Cormier N, Cornwall GA. Alteration in the processing of the ACRBP/sp32 protein and sperm head/acrosome malformations in proprotein convertase 4 (PCSK4) null mice. *Mol Hum Reprod* 2012;**18**:298–307.

Terashima M, Barbour S, Ren J, Yu W, Han Y, Muegge K. Effect of high fat diet on paternal sperm histone distribution and male offspring liver gene expression. *Epigenetics* 2015;**10**:161–171.

Thacker S, Yadav SP, Sharma RK, Kashou A, Willard B, Zhang D, Agarwal A. Evaluation of sperm proteins in infertile men: A proteomic approach. *Fertil Steril* 2011;**95**:2745–2748.

Torabi F, Bogle OA, Estanyol JM, Oliva R, Miller D. Zona pellucida-binding protein 2 (ZPBP2) and several proteins containing BX7B motifs in human sperm may have hyaluronic acid binding or recognition properties. *Mol Hum Reprod* 2017;**23**:803–816.

Torra-Massana M, Cornet-Bartolomé D, Barragán M, Durban M, Ferrer-Vaquer A, Zambelli F, Rodriguez A, Oliva R, Vassena R. Novel phospholipase C zeta 1 mutations associated with fertilization failures after ICSI. *Hum Reprod* 2019;**34**:1494–1504.

Torregrosa N, Domínguez-Fandos D, Camejo MI, Shirley CR, Meistrich ML, Ballescà JL, Oliva R. Protamine 2 precursors, protamine 1/protamine 2 ratio, DNA integrity and other sperm parameters in infertile patients. *Hum Reprod* 2006;**21**:2084–2089.

Trebichalská Z, Holubcová Z. Perfect date—the review of current research into molecular bases of mammalian fertilization. *J Assist Reprod Genet* 2020;**37**:243–256.

Turner S, Wong HP, Rai J, Hartshorne GM. Telomere lengths in human oocytes,

cleavage stage embryos and blastocysts. *Mol Hum Reprod* 2010;**16**:685–694.

Urizar-Arenaza I, Osinalde N, Akimov V, Puglia M, Candenas L, Pinto FM, Muñoa-Hoyos I, Gianzo M, Matorras R, Irazusta J, *et al.* Phosphoproteomic and functional analyses reveal sperm-specific protein changes downstream of kappa opioid receptor in human spermatozoa. *Mol Cell Proteomics* 2019;**18**:S118–S131.

Vågerö D, Pinger PR, Aronsson V, van den Berg GJ. Paternal grandfather's access to food predicts all-cause and cancer mortality in grandsons. *Nat Commun* 2018;**9**:5124.

Van Der Heijden GW, Ramos L, Baart EB, Van Den Berg IM, Derijck AAHA, Van Der Vlag J, Martini E, De Boer P. Sperm-derived histones contribute to zygotic chromatin in humans. *BMC Dev Biol* 2008;**8**:34.

Vandenbrouck Y, Lane L, Carapito C, Duek P, Rondel K, Bruley C, Macron C, Gonzalez de Peredo A, Couté Y, Chaoui K, *et al.* Looking for Missing Proteins in the Proteome of Human Spermatozoa: An Update. *J Proteome Res* 2016;**15**:3998–4019.

Vander Borght M, Wyns C. Fertility and infertility: Definition and epidemiology. *Clin Biochem* 2018;**62**:2–10.

Vassena R, Boué S, González-Roca E, Aran B, Auer H, Veiga A, Belmonte JCI. Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. *Development* 2011;**138**:3699–3709.

Vavouri T, Lehner B. Chromatin Organization in Sperm May Be the Major Functional Consequence of Base Composition Variation in the Human Genome. In Lieb JD, editor. *PLoS Genet* 2011;**7**:e1002036.

Vieweg M, Dvorakova-Hortova K, Dudkova B, Waliszewski P, Otte M, Oels B, Hajimohammad A, Turley H, Schorsch M, Schuppe HC, *et al.* Methylation analysis of histone h4k12ac-associated promoters in sperm of healthy donors and subfertile patients. *Clin Epigenetics* 2015;**7**:31.

Vigodner M, Shrivastava V, Gutstein LE, Schneider J, Nieves E, Goldstein M, Feliciano M, Callaway M. Localization and identification of sumoylated proteins in human sperm: Excessive sumoylation is a marker of defective spermatozoa. *Hum Reprod* 2013;**28**:210–223.

Walser CB, Lipshitz HD. Transcript clearance during the maternal-to-zygotic transition. *Curr Opin Genet Dev* 2011;**21**:431–443.

Wang G, Guo Y, Zhou T, Shi X, Yu J, Yang Y, Wu Y, Wang J, Liu M, Chen X, *et al.* In-depth proteomic analysis of the human sperm reveals complex protein compositions. *J Proteomics* 2013a;**79**:114–122.

Wang G, Wu Y, Zhou T, Guo Y, Zheng B, Wang J, Bi Y, Liu F, Zhou Z, Guo X, *et al.*

Mapping of the N-linked glycoproteome of human spermatozoa. *J Proteome Res* 2013b;**12**:5750–5759.

Wang J, Qi L, Huang S, Zhou T, Guo Y, Wang G, Guo X, Zhou Z, Sha J. Quantitative phosphoproteomics analysis reveals a key role of Insulin Growth Factor 1 Receptor (IGF1R) tyrosine kinase in human sperm capacitation. *Mol Cell Proteomics* 2015;**14**:1104–1112.

Wang XM, Xiang Z, Fu Y, Wu HL, Zhu WB, Fan LQ. Comparative Proteomics Reveal the Association between SPANX Proteins and Clinical Outcomes of Artificial Insemination with Donor Sperm. *Sci Rep* 2018;**8**:6850.

Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat Genet* 2007;**39**:380–385.

Wang Y, Wan J, Ling X, Liu M, Zhou T. The human sperm proteome 2.0: An integrated resource for studying sperm functions at the level of posttranslational modification. *Proteomics* 2016;**16**:2597–2601.

Wang Y, Wang Z. Efficient backsplicing produces translatable circular mRNAs. *RNA* 2015;**21**:172–179.

Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat Rev Genet* 2009;**10**:57–63.

Wei Y, Su J, Liu H, Lv J, Wang F, Yan H, Wen Y, Liu H, Wu Q, Zhang Y. Metaimprint: An information repository of mammalian imprinted genes. *Dev* 2014;**141**:2516–2523.

Williams HL, Mansell S, Alasmari W, Brown SG, Wilson SM, Sutton KA, Miller MR, Lishko P V., Barratt CLR, Publicover SJ, *et al.* Specific loss of CatSper function is sufficient to compromise fertilizing capacity of human spermatozoa. *Hum Reprod* 2015;**30**:2737–2746.

World Health Organization. *Laboratory manual for the examination and processing of human semen*. *Cambridge Cambridge Univ Press* 2010;

Wozniak KL, Carlson AE. Ion channels and signaling pathways used in the fast polyspermy block. *Mol Reprod Dev* 2020;**87**:350–357.

Wu JY, Ribar TJ, Cummings DE, Burton KA, McKnight GS, Means AR. Spermiogenesis and exchange of basic nuclear proteins are impaired in male germ cells lacking Camk4. *Nat Genet* 2000;**25**:448–452.

Wu L, Sampson NS. Fucose, mannose, and β-N-acetylglucosamine glycopolymers initiate the mouse sperm acrosome reaction through convergent signaling pathways. *ACS Chem Biol* 2014;**9**:468–475.

Wu PH, Fu Y, Cecchini K, Özata DM, Arif A, Yu T, Colpan C, Gainetdinov I, Weng Z, Zamore PD. The evolutionarily conserved piRNA-producing locus pi6 is required for male mouse fertility. *Nat Genet* 2020;**52**:728–739.

Xavier MJ, Roman SD, Aitken RJ, Nixon B. Transgenerational inheritance: How impacts to the epigenetic and genetic information of parents affect offspring health. *Hum Reprod Update* 2019;**25**:519–541.

Xu B, Hao Z, Jha KN, Zhang Z, Urekar C, Digilio L, Pulido S, Strauss JF, Flickinger CJ, Herr JC. TSKS concentrates in spermatid centrioles during flagellogenesis. *Dev Biol* 2008;**319**:201–210.

Xu H, Wang X, Wang Z, Li J, Xu Z, Miao M, Chen G, Lei X, Wu J, Shi H, *et al.* MicroRNA expression profile analysis in sperm reveals hsa-mir-191 as an auspicious omen of in vitro fertilization. *BMC Genomics* 2020;**21**:165.

Xu W, Hu H, Wang Z, Chen X, Yang F, Zhu Z, Fang P, Dai J, Wang L, Shi H, *et al.* Proteomic characteristics of spermatozoa in normozoospermic patients with infertility. *J Proteomics* 2012;**75**:5426–5436.

Yan N, Xu H, Zhang J, Xu L, Zhang Y, Zhang L, Xu Y, Zhang F. Circular RNA profile indicates circular RNA VRK1 is negatively related with breast cancer stem cells. *Oncotarget* 2017;**8**:95704–95718.

Yan W, Si Y, Slaymaker S, Li J, Zheng H, Young DL, Aslanian A, Saunders L, Verdin E, Charo IF. *Zmynd15* Encodes a Histone Deacetylase-dependent Transcriptional Repressor Essential for Spermiogenesis and Male Fertility. *J Biol Chem* 2010;**285**:31418–31426.

Yanagimachi R, Knobil E, Neill J. Mammalian fertilization, The Physiology of Reproduction. *Raven Press New York* 1994;, p. 189–317.

Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, Jin Y, Yang Y, Chen LL, Wang Y, *et al.* Extensive translation of circular RNAs driven by N 6 -methyladenosine. *Cell Res* 2017;**27**:626–641.

Yoda A, Oda S, Shikano T, Kouchi Z, Awaji T, Shirakawa H, Kinoshita K, Miyazaki S. Ca2+ oscillation-inducing phospholipase C zeta expressed in mouse eggs is accumulated to the pronucleus during egg activation. *Dev Biol* 2004;**268**:245–257.

Yoon SY, Jellerette T, Salicioni AM, Hoi CL, Yoo MS, Coward K, Parrington J, Grow D, Cibelli JB, Visconti PE, *et al.* Human sperm devoid of PLC, zeta 1 fail to induce Ca2+ release and are unable to initiate the first step of embryo development. *J Clin Invest* 2008;**118**:3671–3681.

Yu H, Diao H, Wang C, Lin Y, Yu F, Lu H, Xu W, Li Z, Shi H, Zhao S, *et al.* Acetylproteomic analysis reveals functional implications of lysine acetylation in

human spermatozoa (sperm). *Mol Cell Proteomics* 2015;**14**:1009–1023.

Yuan S, Schuster A, Tang C, Yu T, Ortogero N, Bao J, Zheng H, Yan W. Sperm-borne miRNAs and endo-siRNAs are important for fertilization and preimplantation embryonic development. *Dev* 2016;**143**:635–647.

Yuan S, Tang C, Zhang Y, Wu J, Bao J, Zheng H, Xu C, Yan W. miR-34b/c and miR-449a/b/c are required for spermatogenesis, but not for the first cleavage division in mice. *Biol Open* 2015;**4**:212–223.

Zhang X, Zhang P, Song D, Xiong S, Zhang H, Fu J, Gao F, Chen H, Zeng X. Expression profiles and characteristics of human lncRNA in normal and asthenozoospermia sperm. *Biol Reprod* 2019a;**100**:982–993.

Zhang XO, Wang H Bin, Zhang Y, Lu X, Chen LL, Yang L. Complementary Sequence-Mediated Exon Circularization. *Cell* 2014;**159**:134–147.

Zhang Y, Shi J, Rassoulzadegan M, Tuorto F, Chen Q. Sperm RNA code programmes the metabolic health of offspring. *Nat Rev Endocrinol* 2019b;**15**:489–498.

Zhang Y, Zhang X-O, Chen T, Xiang J-F, Yin Q-F, Xing Y-H, Zhu S, Yang L, Chen L-L. Circular Intronic Long Noncoding RNAs. *Mol Cell* 2013;**51**:792–806.

Zhao C, Huo R, Wang FQ, Lin M, Zhou ZM, Sha JH. Identification of several proteins involved in regulation of sperm motility by proteomic analysis. *Fertil Steril* 2007;**87**:436–438.

Zhao M, Shirley CR, Hayashi S, Marcon L, Mohapatra B, Suganuma R, Behringer RR, Boissonneault G, Yanagimachi R, Melstrich ML. Transition Nuclear Proteins Are Required for Normal Chromatin Condensation and Functional Sperm Development. *Genesis* 2004;**38**:200–213.

Zhao Y, Li Q, Yao C, Wang Z, Zhou Y, Wang Y, Liu L, Wang Y, Wang L, Qiao Z. Characterization and quantification of mRNA transcripts in ejaculated spermatozoa of fertile men by serial analysis of gene expression. *Hum Reprod* 2006;**21**:1583–1590.

Zhou W, De Iuliis GN, Dun MD, Nixon B. Characteristics of the epididymal luminal environment responsible for sperm maturation and storage. *Front Endocrinol (Lausanne)* 2018;**9**:59.

Zhu Y, Wu Y, Jin K, Lu H, Liu F, Guo Y, Yan F, Shi W, Liu Y, Cao X, *et al.* Differential proteomic profiling in human spermatozoa that did or did not result in pregnancy via IVF and AID. *Proteomics - Clin Appl* 2013;**7**:850–858.

7.

ANNEXOS

7.1. Annex 1

Nucleoproteïnes de l'espermatozoide (Histones i Protamines)

A Clinician's Guide to Sperm DNA and Chromatin Damage

Capítol de Llibre

Sperm Nucleoproteins (Histones and Protamines)
Barrachina F., Soler-Ventura A., Oliva R., Jodar M. (2018) In: Zini A., Agarwal A. (eds) A Clinician's Guide to Sperm DNA and Chromatin Damage. Springer, Cham. https://doi.org/10.1007/978-3-319-71815-6_2

Chapter 2
Sperm Nucleoproteins (Histones and Protamines)

Ferran Barrachina, Ada Soler-Ventura, Rafael Oliva, and Meritxell Jodar

2.1 Introduction

Diploid spermatogonial stem cells differentiate into haploid spermatozoa by an accurately controlled process termed spermatogenesis (Fig. 2.1). Spermatogenesis comprises three distinct phases: a mitotic proliferation phase, a meiotic phase, and the differentiation/maturation phase also known as spermiogenesis [1]. During this last phase, the round spermatids undergo significant nuclear, morphological, and cytoplasmic changes to end up becoming motile, haploid, and highly condensed spermatozoa.

One of the most remarkable features of spermatogenesis is the chromatin dynamics along the different phases (Fig. 2.1) [2–5]. Similarly to somatic cells, the DNA in differentiating spermatogonia is packaged by nucleosomes. Spermatogonia replicate by mitosis to ensure the maintenance of germinal stem cell population. However, certain spermatogonia will enter into meiosis to halve its chromosome content and give rise to haploid germ cells. In the prophase of the first meiotic division, the homologous chromosome recombination occurs. One prerequisite for the homologous recombination is the introduction of DNA double-strand breaks (DSB) and its subsequent repair. A high number of DSB are induced in meiotic cells, and only few of them will be resolved as chromosome crossovers, therefore

F. Barrachina • A. Soler-Ventura • R. Oliva (✉) • M. Jodar (✉)
Molecular Biology of Reproduction and Development Research Group, Department of Biomedicine, Faculty of Medicine, and Biochemistry and Molecular Genetics Service, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, and Hospital Clínic, Casanova 143, 08036 Barcelona, Spain
e-mail: roliva@ub.edu; m.jodar.bifet@gmail.com

A. Zini, A. Agarwal (eds.), *A Clinician's Guide to Sperm DNA and Chromatin Damage*, https://doi.org/10.1007/978-3-319-71815-6_2

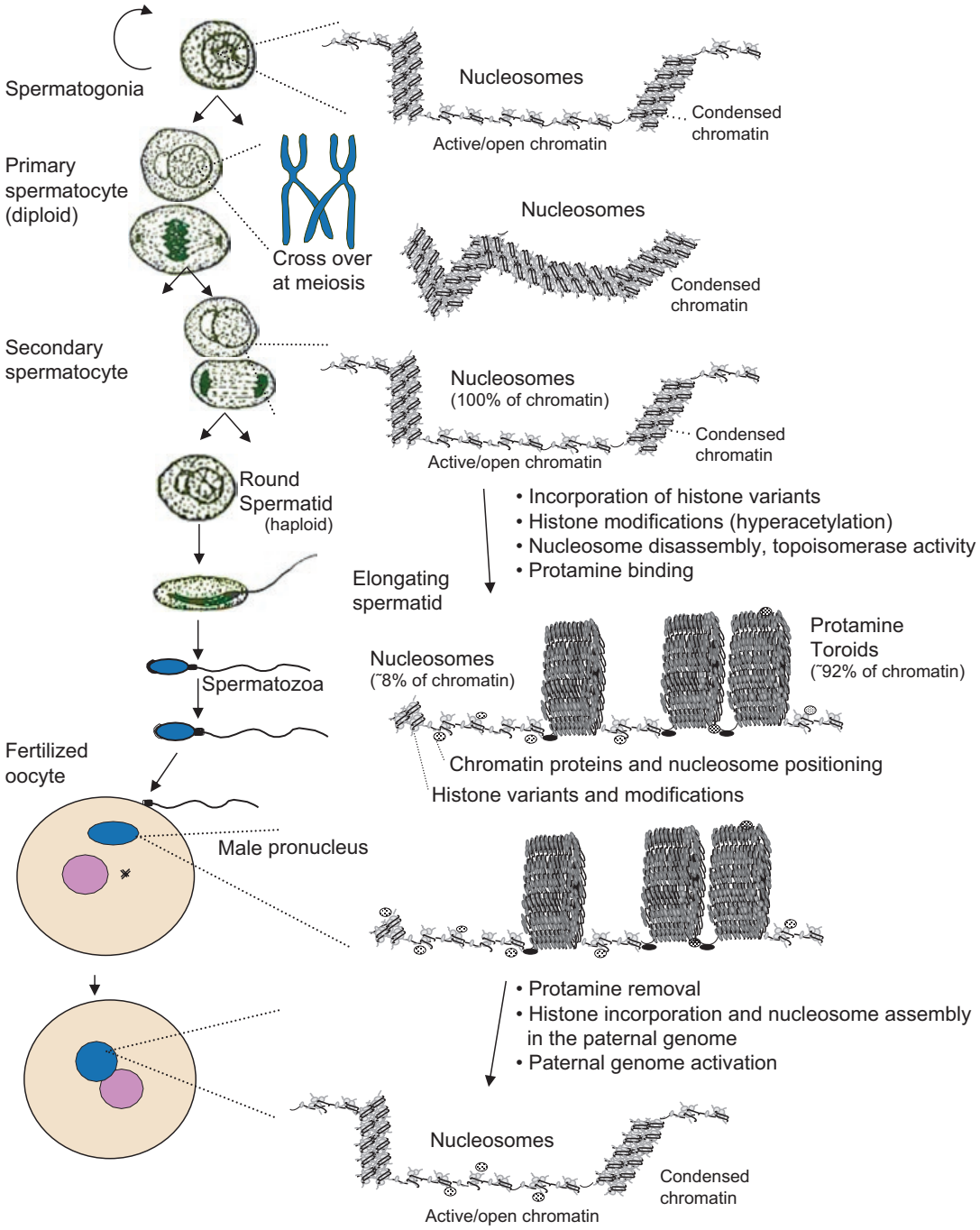


Fig. 2.1 Cellular and chromatin changes during spermatogenesis and fertilization. The main cellular changes (left) are represented together with the concomitant main chromatin changes (right). Spermatogonia replicate and differentiate into primary spermatocytes, which undergo crossing over at meiosis and genetic recombination, and give rise to the secondary spermatocytes after division. Secondary spermatocytes will then divide and give rise to the haploid round spermatids. The round spermatids possess a chromatin structure similar to that of the preceding cells and somatic cells formed by nucleosomes. However, a differentiation process called spermiogenesis is then initiated where the nucleosomal chromatin structure is disassembled and replaced by a highly compact nucleoprotamine complex. The disassembly of nucleosomes changes the superhelicity of the DNA and requires the action of topoisomerases. In the human sperm, about 92% of the chromatin DNA is condensed by protamines forming highly compact toroidal structures each packaging about 50 KB of DNA, and about 8% of the chromatin is formed by nucleosomes. The genes and repetitive sequences are specifically distributed in the nucleohistone and nucleoprotamine structure, and this peculiar chromatin structure is transferred to the oocyte at fertilization. After fertilization, the paternal chromatin must undergo the nucleoprotamine disassembly and the de novo assembly of nucleosomes before paternal gene expression starts

ensuring the genetic variability of the resulting germ cells [6]. In the meiotic and postmeiotic germ cells, the canonical histones are replaced sequentially, first by histone variants [7], subsequently, during spermiogenesis, by transition proteins (TNPs), and, finally, by protamines, following a precise and well-established timing (see Chap. 1) [2, 3, 5, 8]. This process results in a dramatic reorganization of the chromatin exchanging the nucleosomal histone-based structure in the diploid spermatogonia to a nuclear structure tightly packaged by protamines in the haploid spermatozoa, with a potential function in the sperm DNA protection [2, 3]. The multistep procedure of histone exchange requires the contribution of histone variants, as well as histone posttranslational modifications (PTMs); chromatin readers, for example, BRDT [9]; and the transient induction of DSB by topoisomerases to probably eliminate DNA supercoils formed during histone removal [10]. Of relevance, topoisomerases or topoisomerase activity also seems to be present in the final sperm chromatin and may be related to sperm DNA integrity (see Chap. 3) [11]. However, the underlying mechanisms of chromatin reorganization in developing spermatozoa are still poorly understood. Although most of the histones are replaced by protamines during spermatogenesis, the human sperm retains approximately 5–15% of its genome packaged by histones [12]. After fertilization, when the sperm nucleus enters into the oocyte cytoplasm, protamines are quickly replaced by maternal histones, although this process is also poorly understood [13]. However, it has been suggested that the sperm chromatin bound to histones could act as an epigenetic signature with a pivotal role during the activation of zygote genome in early embryogenesis, as well as on transgenerational epigenetic inheritance [14–16].

In this chapter, we highlight the most relevant proteins present in mature spermatozoa, the protamines, and histones, including their variants, their PTMs distribution in the sperm chromatin, and their potential correlation with male infertility.

2.2 Nucleoprotamine Complex in Sperm

Protamines are the most abundant sperm nuclear proteins in many species and in human are packing approximately the 85–95% of the paternal DNA [2, 3, 17–19]. Protamines are small basic proteins rich in positively charged arginine residues, allowing the formation of a highly condensed complex with the negatively charged paternal DNA. Additionally, protamines are rich in cysteine residues, which allow the formation of disulfide bonds and zinc bridges among intra- and inter-protamine molecules resulting in the compact toroidal nucleoprotamine complex [20, 21]. In mammals, two types of protamines have been described, the protamine 1 (P1) and the protamine 2 (P2) family. All mammal species harbor P1 in spermatozoa, but the P2 family, composed by the P2, P3, and P4 components, is solely expressed by some mammal species, such as humans and mice [18, 22]. Typically, the genes encoding protamines (*PRM1* and *PRM2*) are clustered together. In human, the protamine gene cluster is located in chromosome 16 together with the transition nuclear protein 2 (*TNP2*) gene [23]. Whereas P1 is synthesized as a mature form, P2 family is generated from the proteolysis of the protamine 2 precursor resulting in the

different components of P2 family (P2, P3, P4), which differ among them only by one to four amino acid residues on the N-terminal extension, being the P2 the most abundant [17, 18].

Although several hypotheses of the P1 and P2 family functions have been proposed [2, 3, 18], the most accepted protamine functions are:

(i) To tightly package the paternal genome in a more compact and hydrodynamic nucleus required for a proper sperm motility
(ii) To protect the paternal genome from exogenous or endogenous mutagens or nucleases potentially present in the male and/or female tracts
(iii) To compete with and remove transcriptional factors and other nuclear proteins from the spermatid chromatin, leaving the paternal genome in a "blank state" so that the paternal genome could be reprogrammed by the oocyte
(iv) To be involved in the imprinting of the paternal genome during spermatogenesis and to confer new epigenetic marks in certain areas of the sperm genome, leading to gene reactivation or repression in the first steps of early embryo development [3, 18]

2.2.1 Protamine Post-translational Modifications

In contrast to the well-known roles of histone PTMs, such as acetylation, methylation, and phosphorylation (see Sect. 3.2), relatively little is known about protamine PTMs. The most well-studied protamine PTM has been phosphorylation (Fig. 2.1, Table 2.1) [2, 3]. Protamines are quickly phosphorylated after their synthesis in elongated spermatids, as a requisite for the proper protamine binding to sperm DNA [17]. However, after the protamine-DNA binding, protamines are extensively dephosphorylated except in some residues whose phosphorylation can still be observed in the mature sperm (Fig. 2.2, Table 2.1) [24, 27–29]. Another type of protamine PTM is the differential processing of protamine 2 precursors. In fact, protamine 2 is synthesized as a long precursor protein which is then proteolytically processed to give rise to the mature P2, P3, and P4 components [30]. More recently, the use of mass spectrometry has allowed to identify additional PTMs in both protamines, suggesting the existence of a protamine code similar to the histone code [31, 32] that could be relevant for zygote epigenetic reprogramming [26, 33, 34]. In mature human sperm, the analysis of the extracted intact protamines by mass spectrometry has enabled to identify mono-, di-, and tri-phosphorylations, di-acetylations, and a mono-methylation for P1 [25]. Using the same strategy, only the intact P3 component could be identified from the P2 family with two potential PTMs (one acetylation and one methylation) [25]. However, further studies are required in humans including the amino acid sequencing by mass spectrometry in order to identify new protamine PTMs and localize the modified residues, as has been recently described in mouse (Fig. 2.2, Table 2.1) [26].

Table 2.1 Posttranslational modifications (PTMs) detected in human and mouse protamine amino acid sequences

Specie	Protamine	Amino acid residue	Post-translational modification	Methodology	Reference
Human	Protamine 1	S8	Phosphorylation	Electrospray mass spectrometry	Chirat et al. [24]
				Phosphoserine conversion and protein sequencing	Pirhonen et al. [29]
		S10	Phosphorylation	Electrospray mass spectrometry	Chirat et al. [24]
				Phosphoserine conversion and protein sequencing	Pirhonen et al. [29]
		S28	Phosphorylation	Phosphoserine conversion and protein sequencing	Pirhonen et al. [29]
		ND	Phosphorylation	Mass spectrometry	Castillo et al. [25]
		ND	Acetylation	Mass spectrometry	Castillo et al. [25]
	Protamine 2	S50	Phosphorylation	Phosphoserine conversion and protein sequencing	Pirhonen et al. [29]
		S58	Phosphorylation	Electrospray mass spectrometry	Chirat et al. [24]
				Phosphoserine conversion and protein sequencing	Pirhonen et al. [29]
		S72	Phosphorylation	Phosphoserine conversion and protein sequencing	Pirhonen et al. [29]
Mouse	Protamine 1	S8	Phosphorylation	Mass spectrometry	Brunner et al. [26]
		S42	Phosphorylation	Mass spectrometry	Brunner et al. [26]
		S42	Acetylation	Mass spectrometry	Brunner et al. [26]
		T44	Phosphorylation	Mass spectrometry	Brunner et al. [26]
		K49	Methylation	Mass spectrometry	Brunner et al. [26]
		K49	Acetylation	Mass spectrometry	Brunner et al. [26]
		N-terminal	Acetylation	Mass spectrometry	Brunner et al. [26]
	Protamine 2	S55	Phosphorylation	Mass spectrometry	Brunner et al. [26]
		S55	Acetylation	Mass spectrometry	Brunner et al. [26]
		K57	Acetylation	Mass spectrometry	Brunner et al. [26]
		K64	Acetylation	Mass spectrometry	Brunner et al. [26]

The table shows the PTMs identified in human and mouse protamine 1 or protamine 2 amino acid residues and the methodology performed
ND Not determined

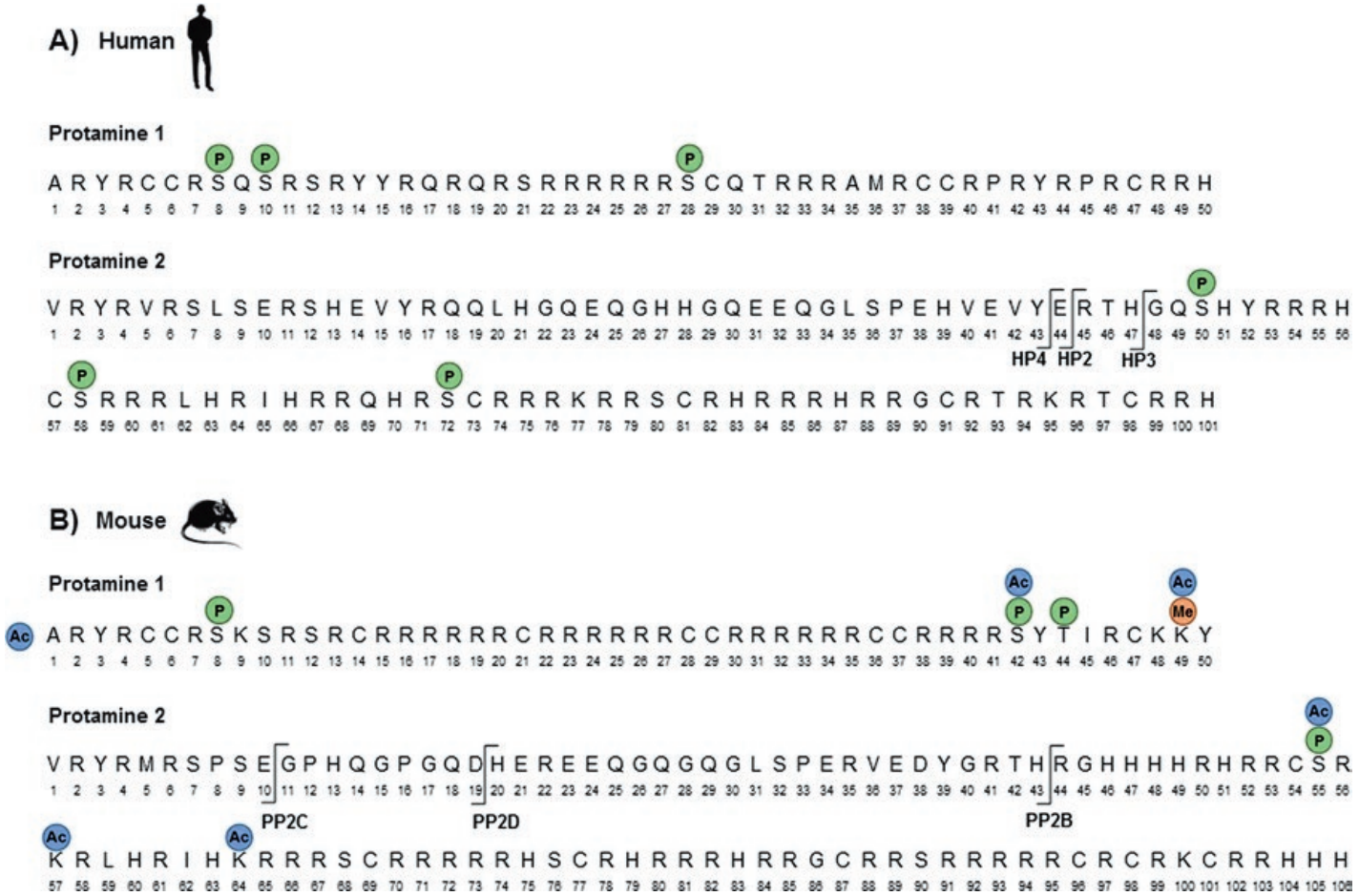


Fig. 2.2 Protamine amino acid sequences and posttranslational modifications (PTMs) detected in human and mouse spermatozoa. PTMs are indicated by Ac for acetylation, Me for methylation, and P for phosphorylation. Residues are numbered after the cleaved methionine. (**a**) Human protamine 1 and protamine 2 amino acid sequences with the described PTMs. The human HP2, HP3, and HP4 cleavage sites are indicated in the sequence. (**b**) Mouse protamine 1 and protamine 2 amino acid sequence with the described PTMs. The mouse PP2B, PP2C, and PP2D cleavage sites are indicated in the sequence

2.2.2 *Protamine Alterations in Infertile Patients*

The relative ratio of the abundance of the protamine 1 (P1) and the protamine 2 (P2) has been widely studied as a measure of sperm chromatin maturity and normality/abnormality [18]. A prospective study in the general population proposed the presence of a wide range for P1/P2 ratio that can oscillate between 0.5 and 1.5 [35]. However, from a reproductive view, several groups proposed a P1/P2 ratio around 1 (0.8–1.2) for fertile males [36, 37]. An altered P1/P2 ratio (below 0.8 or above 1.2) has been correlated among seminal parameter alterations, DNA damage, and low success rate of assisted reproduction techniques (Table 2.2). A recent meta-analysis comparing infertile and fertile patients, including data from nine different studies, has demonstrated a significantly increased P1/P2 ratio in subfertile patients [57].

P2 deregulation occurs more frequently than P1 deregulation, indicating that a P2 deregulation is normally responsible for the P1/P2 ratio alteration. Lower sperm count and sperm motility and/or abnormal sperm morphology have been correlated with abnormal P1/P2 ratio (Table 2.2) [38–45, 47, 48, 58, 59]. Furthermore, some studies have shown that the total absence of P2 or the incomplete processing of the P2 precursors reflected by a decreased pre-P2/P2 ratio is also linked to a lower sperm count, a lower sperm motility, and an abnormal sperm morphology (Table 2.2) [41, 44–46]. Additionally, an altered P1/P2 ratio or a decreased pre-P2/P2 ratio was also linked to an increased sperm DNA damage or to an augmented reactive oxygen species levels (Table 2.2) [41, 49–54, 58]. These studies suggest that an altered P1/P2 ratio results in a sperm DNA more accessible to nuclease activity and, therefore, DNA damage increases. A correct protamination, as a measure of a correct P1/P2 ratio, could be crucial for the DNA protection [49]. However, the meta-analysis performed by Ni et al. could not establish an association between an altered P1/P2 ratio and DNA damage [57]. Otherwise, several studies have also correlated an altered P1/P2 ratio with a low fertilization rate, a low implantation rate, a low embryo quality score, and a low pregnancy outcome using in vitro fertilization (IVF) with or without intracytoplasmic sperm injection (ICSI) treatments (Table 2.2) [40, 43, 45, 52, 55, 59]. Likewise, a decreased pre-P2/P2 ratio has been correlated to a low implantation rate and a poor pregnancy outcome [45], and the total absence of P2 has been correlated to a low sperm penetration ability in IVF [46]. Taken together, these results suggest that protamine deregulation could be involved in fertilization and early embryo development processes. Other parameters such as men's age and smoking have been proposed to alter the protamine P1/P2 ratio (Table 2.2) [38, 39, 54]. In addition, it has been reported that a mutation in the *PRM1* gene promoter (−191AA genotype) causes an increased P1/P2 ratio suggesting that genetic mutations could be the cause of a defective protamination [56]. All these studies suggest that a correct P1/P2 ratio is important for men's fertility and for proper embryo development.

Table 2.2 Altered P1/P2 ratio in infertile patients

Study	P1/P2 ratio	Outcome
Correlation with seminal parameters		
Simon et al. [38]	Altered P1/P2	Lower sperm count; lower semen volume
Hamad et al. [39]	Increased P1/P2	Lower sperm count; lower sperm vitality
Aoki et al. [40]	Altered P1/P2	Lower sperm count; lower sperm motility; abnormal sperm morphology
Torregrosa et al. [41]	Decreased pre-P2/P2	Lower sperm count; lower sperm motility; abnormal sperm morphology
Aoki et al. [58]	Altered P1/P2	Lower sperm count; lower sperm motility; abnormal sperm morphology
Mengual et al. [42]	Increased P1/P2	Lower sperm count
Khara et al. [43]	Altered P1/P2	Lower sperm count; lower sperm motility; abnormal sperm morphology
de Yebra et al. [44]	Altered P1/P2	Lower sperm count
	No P2	Lower sperm count; lower sperm motility; abnormal sperm morphology
De Mateo et al. [45]	Decreased P1/P2	Lower sperm motility
	Decreased pre-P2/P2	Lower sperm count; lower sperm motility
Aoki et al. [59]	Altered P1/P2	Lower sperm count; lower sperm motility; abnormal sperm morphology
	Decreased P1/P2	Abnormal sperm head morphology
Carrell and Liu [46]	No P2	Lower sperm motility; abnormal sperm morphology
Bach et al. [47]	Altered P1/P2	Altered seminal parameters
Lescoat et al. [48]	Altered P1/P2	Altered seminal parameters
Correlation with DNA damage		
Ribas-Maynou et al. [49]	Increased P1/P2	Increased DNA damage (SCD assay)
García-Peiró et al. [50]	Increased P1/P2	Increased DNA damage (SCD assay)
Castillo et al. [51]	Decreased P1/P2	Increased DNA damage (alkaline comet assay)
Simon et al. [52]	Increased P1/P2	Increased DNA damage (alkaline comet assay)
Aoki et al. [53]	Altered P1/P2	Increased DNA damage (TUNEL assay)
Torregrosa et al. [41]	Decreased pre-P2/P2	Increased DNA damage (TUNEL assay)
Aoki et al. [58]	Decreased P1/P2	Increased DNA damage (SCSA assay)
Hammadeh et al. [54]	Increased P1/P2	Increased reactive oxygen species (ELISA assay)
Correlation with assisted reproduction techniques		
Simon et al. [52]	Decreased P1/P2	Low fertilization rate (IVF)
De Mateo et al. [45]	Decreased P1/P2	Low fertilization rate (IVF); low implantation rate (IVF and/or ICSI); low pregnancy outcome (IVF and/or ICSI)
	Decreased pre-P2/P2	Low implantation rate (IVF and/or ICSI); low pregnancy outcome (IVF and/or ICSI)
Aoki et al. [40]	Altered P1/P2	Low fertilization rate (IVF)
	Decreased P1/P2	Low chemical-pregnancy and clinical-pregnancy rates (IVF and/or ICSI)

(continued)

Table 2.2 (continued)

Study	P1/P2 ratio	Outcome
Aoki et al. [59]	Decreased P1/P2	Low fertilization rate (IVF and ICSI)
Nasr-Esfahani et al. [55]	Increased P1/P2	Low fertilization rate (ICSI); low embryo quality score in day 3 (ICSI)
Khara et al. [43]	Altered P1/P2	Low fertilization rate (IVF)
Carrell and Liu [46]	No P2	Low sperm penetration ability (IVF)
Correlation with other parameters		
Simon et al. [38]	Altered P1/P2	Men's age
Hamad et al. [39]	Increased P1/P2	Smokers
Hammadeh et al. [54]	Increased P1/P2	Smokers
	Decreased P2	Smokers
Jodar et al. [56]	Increased P1/P2	Mutation in the PRM1 gene promoter (−191AA genotype)

Correlation of protamine P1/P2 ratio with seminal parameters, DNA damage, assisted reproduction techniques outcome, and other parameters
SCD sperm chromatin dispersion, *TUNEL* terminal deoxynucleotidyl transferase dUTP nick end labeling, *SCSA* sperm chromatin structure assay, *ELISA* enzyme-linked immunosorbent assay, *IVF* in vitro fertilization, *ICSI* intracytoplasmic sperm injection

2.3 Nucleohistone Complex in Sperm

As mentioned before, the human spermatozoon retains approximately a 5–15% of its chromatin packaged in nucleosomes [18]. The nucleosome structure in sperm seems to be similar to that from somatic cells and consists on 147 base pairs of DNA wrapped around an octameric histone core including two of each H2A, H2B, H3, and H4 histones [60, 61]. Adjacent nucleosomes are interconnected by a linker DNA that can be up to 80 bp long. Members of histone H1 family (linker histones) are situated at the site of DNA entry and exit from the core particle binding around 20 nt of linker DNA. Apart from acting as a linker, histone H1 plays an important role in the chromatin folding modulation. The final result is a constrained DNA that approximately achieves a fivefold compaction. Despite the high degree of compaction that nucleosomes confer, histone-packaged chromatin sperm is more open and dynamic than the protamine-packaged chromatin and could be modulated and regulated by the incorporation of histone variants [7, 62], histone PTMs [31], and nuclear factors that modulate the DNA and histone interactions [3, 63].

2.3.1 Histone Variants

During spermatogenesis, some canonical histones are replaced by histone variants, and a subset of those remains in the nucleus of mature spermatozoa. Several histone variants have been identified in mature sperm by mass spectrometry, including the histone H4, which is less diversified compared with the most diverse H2A and H2B histones (Table 2.3). Despite the fact that histone variants have only small changes

Table 2.3 Human sperm histone variants identified in mature sperm

Protein name	Gene name	Localization	♂KO effect on reproduction and embryogenesis	References
Histone H1 family				
Histone H1t	HIST1H1T	Testis	Normal phenotype	Lin et al. [64], Fantz et al. [65]
Testis-specific H1 histone (H1t2)	H1FNT	Testis	Oligozoospermia, asthenozoospermia, teratozoospermia, abnormal spermiogenesis, reduced male fertility, and impaired fertilization	Martianov et al. [66], Tanaka et al. [67]
Histone H1x	H1FX	All tissues	ND	–
Histone H1.2	HIST1H1C	All tissues	Normal phenotype	Fan et al. [68]
Histone H1.3	HIST1H1D	All tissues	Normal phenotype	
Histone H1.4	HIST1H1E	All tissues	Normal phenotype	
Histone H1.5	HIST1H1B	All tissues	ND	–
Histone H2 family				
Histone H2A type 1	HIST1H2AG	Testis	ND	–
Histone H2A type 1-A (TH2A)	HIST1H2AA	Testis	ND	–
Histone H2B type 1-A (TH2B)	HIST1H2BA	Testis	ND	–
Histone H2A-Bbd type 1	H2AFB1	Testis	ND	–
Histone H2A-Bbd type 2/3	H2AFB2	Testis	ND	–
Histone H2A type 1-B/E	HIST1H2AB	Enriched in testis	ND	–
Histone H2A type 1-H	HIST1H2AH	Enriched in testis	ND	–
Histone H2B type 1-B	HIST1H2BB	Enriched in testis	ND	–
Histone H2B type 1-J	HIST1H2BJ	Enriched in testis	ND	–
Core histone macro-H2A.1 (mH2A1)	H2AFY	All tissues	Normal phenotype	Changolkar et al. [69], Boulard et al. [70]
Histone H2A type 1-C	HIST1H2AC	All tissues	ND	–
Histone H2A type 2-A	HIST2H2AA3	All tissues	NP	–
Histone H2A type 2-C	HIST2H2AC	All tissues	ND	–
Histone H2A.V	H2AFV	All tissues	ND	–
Histone H2AX	H2AFX	All tissues	Seminiferous tubules reduced diameter, small testes, male meiosis arrest, and male infertility	Celeste et al. [71]

(continued)

Table 2.3 (continued)

Protein name	Gene name	Localization	♂KO effect on reproduction and embryogenesis	References
Histone H2A.Z	H2AFZ	All tissues	Not viable	Faast et al. [72]
Histone H2B type 1-C/E/F/G/I	HIST1H2BC	All tissues	ND	–
Histone H2B type 1-D	HIST1H2BD	All tissues	NP	–
Histone H2B type 1-H	HIST1H2BH	All tissues	ND	–
Histone H2B type 1-K	HIST1H2BK	All tissues	ND	–
Histone H2B type 1-L	HIST1H2BL	All tissues	ND	–
Histone H2B type 1-M	HIST1H2BM	All tissues	ND	–
Histone H2B type 1-N	HIST1H2BN	All tissues	ND	–
Histone H2B type 1-O	HIST1H2BO	All tissues	NP	–
Histone H2B type 2-E	HIST2H2BE	All tissues	KO not fertility related	Santoro et al. [73]
Histone H2B type 2-F	HIST2H2BF	All tissues	NP	–
Histone H2B type 3-B	HIST3H2BB	All tissues	ND	–
Histone H2B type F-S	H2BFS	All tissues	NP	–
Histone H2A type 1-D	HIST1H2AD	–	ND	–
Histone H2A type 2-B	HIST2H2AB	–	ND	–
Histone H3 family				
Histone H3.1	HIST1H3A	Testis	ND	–
Histone H3.1 t (H3t)	HIST3H3	Testis	NP	–
Histone H3.3C	H3F3C	Testis	NP	–
Histone H3.2	HIST2H3A	All tissues	ND	–
Histone H3.3	H3F3A	All tissues	Reduced male fertility	Tang et al. [74]
Histone H3-like centromeric protein A (CENP-A)	CENPA	All tissues	Not viable	Howman et al. [75], Kalitsis et al. [76]
Histone H4 family				
Histone H4	HIST1H4L	All tissues	ND	–
Histone H4	HIST1H4A	All tissues	ND	–
Histone H4-like protein type G	HIST1H4G	ND	NP	–

Integrative table of the human sperm histone families combining the protein/gene name, GTEx localization, and the knockout effect on reproduction/embryogenesis using Mouse Genome Informatics database
ND no data, *NP* not present in mouse

in their primary structure compared with the canonical histones, those little differences can lead to major changes in the nucleosome structure, stability, and function [62]. The destabilization of DNA-protein interaction by incorporation of histone variants during spermatogenesis allows the transition from the nucleohistone complex to the nucleoprotamine complex [7, 18, 62, 77].

Although there are histone variants widely expressed in all tissues, there are some testis-specific variants that are essentially expressed in spermatocytes [78]. Targeting the individual histone variants in mouse models (knockouts) has revealed which histone variants are crucial for male fertility and reproduction (Table 2.3). Unfortunately, there is a lack of information about a set of histone variants that are not present or have not been detected in mouse (NP) or the corresponding knockout model has not been generated yet (ND). In addition, it is not possible to assess the effect on male reproduction of some histone variants because the knockouts have resulted in embryonic lethality [72, 75, 76], pointing out the need to generate conditional knockout models to assess their importance in testes function (Table 2.3).

Knockouts of some histone variants display a normal phenotype without negative impact on fertility, for example, histone H1t, mH2A1, H2B type 2-E, H1.2, H1.3, and H1.4, suggesting that they are not essential for male fertility (Table 2.3) [64, 65, 68–70, 73]. However, it could be expected that different testis-specific histone variants should have a major importance for proper fertility. As observed in Table 2.3, the knockout models of some testis-specific histones or widely expressed histone variants seem to result in reproductive failure. This is the case of a testis-specific histone, the H1t2, and the widely expressed histones H2AX and H3.3. Each knockout of these three different histone variants displayed male infertility although due to different reasons. For example, H1t2 knockout displays an abnormal spermatogenesis, sperm defects, and impaired fertilization, because this histone is necessary for DNA condensation and nuclear modulation during the last steps of spermatogenesis [66, 67]. In contrast, the disruption of H2AX and H3.3 in mice results in male meiosis arrest, since H2AX is crucial for meiosis because it facilitates the repair of induced DSBs [62, 71] and H3.3 is essential for chromosome segregation that takes place during meiosis (Table 2.3) [74].

2.3.2 Histone Post-translational Modifications

The early events during the transition of histones to protamines throughout spermiogenesis involve the incorporation of histone variants and histone PTMs, which enable the chromatin remodeling and trigger the protamination. Both histones and histone variants are modified by different PTMs [79]. The most known histone PTMs are acetylation, methylation, phosphorylation, ubiquitination, sumoylation, and ribosylation, among other forms [79]. The different combinatorial patterns of the huge number of histone PTMs create a complex histone code that contributes to chromatin organization and dynamics, as well as to gene expression [7, 60]. For example, the massive increase of histone acetylation is one of the first signs that the

protamine replacement during spermiogenesis will start [3, 18, 80–82]. Histone hyperacetylation relaxes the chromatin and decreases the affinity of the sperm histones to the DNA, allowing the removal and replacement of histones, firstly, by TNPs and, finally, by protamines [3, 18, 63, 83]. Actually, histone H4 hyperacetylation in elongating spermatids is a prerequisite for the histone-to-protamine replacement [84], and an aberrant H4 hyperacetylation pattern results in impaired spermatogenesis [3, 18, 81, 85]. Apart from histone acetylation, there are other histone modifications throughout spermatogenesis such as methylation, which could also be associated with nucleosome dismantlement and histone eviction. Histone methylation seems to modulate epigenetic signals necessary for spermatogenesis [14, 60]. This is the case of H3K4me, a methylation mark that is necessary to turn spermatogonia into spermatocytes [86, 87], and H3K9me and H3K27me, marks that regulate gene expression during spermatogenesis [88].

Although histone variants and histone PTMs allow chromatin remodeling and most of them are replaced by protamines during late spermatogenesis, some of the modified nucleosomes are not replaced and are retained in the mature sperm. More than 100 histone post-translational modifications have been identified in the remaining histones present in human sperm, including acetylation, methylation, phosphorylation, butyrylation, and crotonylation [31, 89, 90]. Surprisingly, some histone PTMs found in human mature sperm showed a high degree of conservation with mouse sperm, which further supports an evolutionary conserved role of histone PTMs [31]. Those modified paternal histones maintained in the sperm are inherited by the zygote, and they have been proposed to play a role in the epigenetic control of embryogenesis [34]. For example, alterations in the histone methylation (H3K4me2) pattern in mice sperm impair the development and survivability of the offspring, indicating the importance of the sperm epigenome in the health of the progeny [91].

2.3.3 Histone-Bound Sperm Chromatin

Many studies suggest that the 5–15% retained nucleosomes in mature sperm are not randomly distributed through the sperm genome but occupy specific loci [12, 14, 15, 19, 32, 89, 92, 93]. This is supported by recent sperm chromatin high-throughput genome-wide dissection studies indicating that there is a differential distribution of genes and repetitive sequences between nucleohistone and nucleoprotamine complexes.

The first studies using human sperm chromatin fractionation followed by microarrays or high-throughput sequencing concluded that mature sperm histones are associated with DNA enriched at gene regulatory regions and genes involved in developmental processes, including promoters of embryonic transcription factors and signaling pathway proteins, as well as miRNA clusters and imprinted genes (Table 2.4) [19, 89, 92]. In contrast, protamines seemed to be enriched at olfactory receptors genes and ZNF genes [92]. Interestingly, the use of sperm chromatin immunoprecipitation (ChIP) of specific histone PTMs followed either by microarray

Table 2.4 Sperm nucleosomal DNA distribution in healthy men

Study	Methodology	Main outcomes
Arpanahi et al. [92]	Sperm salt extraction and endonuclease digestion or micrococcal nuclease (MNase) digestion followed by a microarray-based genome-wide analysis. Additionally, after digestion, ChIP-chip for acH4 was also used	Endonuclease-sensitive DNA regions are enriched in gene regulatory regions including promoter sequences involved in the development and CTCF-recognized sequences
Hammoud et al. [89]	Sperm MNase digestion followed by either array analysis or high-throughput sequencing. Additionally, ChIP-chip and ChIP-seq for H3K9me3, H3K27me3, H3K4me2/me3, TH2B, and H2A.Z were performed	Sperm nucleosomes are enriched at loci of developmental importance including imprinted gene clusters, miRNA clusters, HOX gene clusters, and promoters of embryo developmental transcription and signaling factors. Histone modifications (H3K4me2/3 and H3K27me3) localize to particular developmental loci
Brykczynska et al. [32]	Sperm MNase digestion followed by mononucleosomal DNA isolation and ChIP for H3K4me2 and H3K27me3 combined with microarray analysis or high-throughput sequencing	Sperm nucleosomes are slightly enriched at TSS. H3K27me3 and H3K4me2 are retained at regulatory sequences in mature human spermatozoa and marks promoters of genes related with spermatogenesis and early embryonic development
Vavouri et al. [15]	Reanalysis of the data from Arpanahi et al. [92], Hammoud et al. [89], and Brykczynska et al. [32]	Nucleosome retention, which is determined by the base composition, occurs in both genic and nongenic regions of the genome. Nucleosomes at GC-rich sequences with high nucleosome affinity are retained at TSSs and at developmental regulatory genes, particularly TSSs of most housekeeping genes. Also, there is a link between nucleosome retention in sperm and DNA unmethylated regions in the early embryo
Samans et al. [93]	Sperm cell fractionation by micrococcal nuclease followed by DNA high-throughput sequencing of the nucleosomal fraction	Sperm chromatin nucleosomes are enriched in certain repetitive DNA elements, as centromere repeats and retrotransposons (LINE1 and SINEs), and the majority of nucleosomal binding sites are enriched in distal intergenic regions. Nucleosome depletion was observed within exons, the majority of promoters, 5′-UTRs, 3′-UTRs, TSS, and TTS. Function of paternally derived nucleosomes in postfertilization processes
Castillo et al. [19]	Sperm chromatin fractionation using salt extraction followed by restriction enzyme digestion or MNase digestion, followed by high-throughput sequencing and proteomic analyses (LC-MS/MS)	Nucleosomal and subnucleosomal DNA regions are highly enriched at gene promoters, CpG island promoters, and linked to genes involved in embryo development

TSS Transcription start site, *TTS* Transcription termination site

(ChIP-chip) or DNA sequencing (ChIP-seq) has revealed that H3K4me2 and H3K4me3 are enriched at developmental promoters expressed in the four- to eight-cell stage embryos, suggesting a potential epigenetic function of those modified sperm histones in early embryogenesis [89]. The specific study of sperm mononucleosomal DNA has shown slight differences, for example, H3K4me2 marks genes involved in spermatogenesis and cellular homeostasis, while H3K27me3 marks developmental regulators and HOX genes [32]. These differences could be attributed to different technical issues in the preparation of the human sperm mononucleosomal DNA in contrast to all nucleohistone complex [32]. In silico analysis from the studies mentioned above revealed that spermatozoal nucleosomes are retained at GC-rich loci and that nucleosome retention in the sperm cell is linked to demethylated DNA in the early embryo [15].

In contrast to the mentioned findings above, one study claimed that retained nucleosomes in sperm are enriched in certain repetitive DNA sequences, such as centromere repeats and retrotransposons (LINE1 and SINE), and the majority of nucleosomal binding sites were enriched at distal intergenic regions [93]. However, these contradictory observations are probably due to technical issues or differences in the computational methodology used [94].

As a summary, there is huge evidence suggesting the existence of a differential distribution between histone-packaged and protamine-packaged sperm chromatin, which is involved in a potential sperm epigenetic signature transferred into the oocyte. The sperm nucleosome enrichment at developmental regulatory genes and gene regulatory sequences suggest that it could regulate the gene expression in early embryogenesis when zygote genome activation occurs and indicate that sperm chromatin is much more complex than it was previously thought.

2.3.4 Histone Alterations and Male Infertility

In contrast to the vast number of studies assessing the potential correlation between protamines (P1/P2 ratio) and male infertility (see Sect. 2.2), very few studies have evaluated sperm histones in infertile patients. Early obervations already indicated that a large proportion of the sperm samples with an altered P1/P2 ratio also had increased levels of histones [3, 18, 44]. Focusing on specific histones, it has been described that γH2AX levels are higher in the sperm of infertile patients than in fertile men, and it has been correlated to an increased number of sperm DSBs [95]. It has also been reported that semen samples from infertile men have a significant higher H2B/(P1+P2) ratio than do fertile men, suggesting that an alteration of H2B/(P1+P2) ratio could reflect an abnormal chromatin structure that results in male infertility [96–98]. Moreover, it has also been found an increased H2B/(P1+P2) ratio in smokers [39], implying a negative effect between smoking cigarettes and male fertility. Finally, a correlation has been found between alterations of a testis-specific histone variant (TH2B) and male fertility, which indicates that TH2B is involved in sperm chromatin compaction and male pronucleus development [99].

Apart from the abovementioned alterations in histone content, the sperm of infertile men has also shown an altered histone localization pattern [100]. The study of these infertile men revealed a randomly distributed pattern of nucleosome retention in the sperm chromatin [100]. This alteration in nucleohistone-bound genome could be attributed to a disrupted chromatin remodeling machinery or due to an improper histone hyperacetylation signaling during the histone exchange by protamines [100]. On a different line of experiments, evidence for a substantial deregulation of histones has been detected in normozoospermic sperm cells from male infertile patients with failed assisted reproduction outcomes after ICSI [101]. Overall, these studies demonstrate the importance of an appropriate distribution of genes in the sperm chromatin structure. Therefore, the potential side effects in the embryo associated to an improper histone retention in the sperm are an aspect that deserves further investigation in the future.

2.4 Concluding Remarks

Protamines have been largely studied and correlated with male infertility, specifically by P1/P2 ratio measurement. Similarly, alterations of specific histones have also been associated with sperm defects. Recent studies support the idea that the distribution of the nucleohistone and nucleoprotamine complexes in the sperm chromatin is not random. The intracytoplasmic sperm injection (ICSI) of mouse round spermatids, that did not complete the histone replacement yet, into mature oocytes, derived in embryos with aberrant patterns of gene expression, thereby suggesting that the paternal chromatin structure is important for the first steps of early embryo development [5]. The complexity of sperm chromatin highlights the need to perform further studies in sperm nucleoproteins content and distribution, including the assessment of their variants and PTMs, in order to clarify the significance of the sperm chromatin in male infertility and early embryo development as well as to shed light into the possible effects across generations. Furthermore, it will be particularly interesting to determine the specific role of the hundreds of chromatin-associated proteins present in the normal sperm chromatin, in addition to histones and protamines, as derived from recent high-throughput proteomic studies [102–104].

Acknowledgments Supported by grants to RO from the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competividad; Instituto de Salut Carlos III, Fondos FEDER, “una manera de hacer Europa,” PI13/00699, P16/00346, CP11/00312, BFU2011–29739), EUGIN-UB Research Excellence Program (EU-REP 2014), Fundación Salud 2000 (SERONO 13–015), and EU-FP7-PEOPLE-2011-ITN289880. FB is granted by the Spanish Ministry of Education, Culture and Sports (Ministerio de Educación, Cultura y Deporte para la Formación de Profesorado Universitario, FPU15). MJ is granted by the Government of Catalonia (Generalitat de Catalunya, pla estratègic de recerca i innovació en salut, PERIS 2016–2020).

References

1. Neto FTL, Bach PV, Najari BB, Li PS, Goldstein M. Spermatogenesis in humans and its affecting factors. Semin Cell Dev Biol. 2016;59:10–26.
2. Balhorn R. The protamine family of sperm nuclear proteins. Genome Biol. 2007;8:227.
3. Oliva R, Dixon GH. Vertebrate protamine genes and the histone-to-protamine replacement reaction. Prog Nucleic Acid Res Mol Biol. 1991;40:25–94.
4. Carrell DT, Aston KI, Oliva R, Emery BR, De Jonge CJ. The "omics" of human male infertility: integrating big data in a systems biology approach. Cell Tissue Res. 2016;363(1):295–312.
5. Rathke C, Baarends WM, Awe S, Renkawitz-Pohl R. Chromatin dynamics during spermiogenesis. Biochim Biophys Acta – Gene Regul Mech. 2014;1839(3):155–68.
6. Youds JL, Boulton SJ. The choice in meiosis – defining the factors that influence crossover or non-crossover formation. J Cell Sci. 2011;124:501–13.
7. Kimmins S, Sassone-Corsi P. Chromatin remodelling and epigenetic features of germ cells. Nature. 2005;434:583–9.
8. Bao J, Bedford MT. Epigenetic regulation of the histone-to-protamine transition during spermiogenesis. Reproduction. 2016;151:R55–70.
9. Goudarzi A, Shiota H, Rousseaux S, Khochbin S. Genome-scale acetylation-dependent histone eviction during spermatogenesis. J Mol Biol. 2014;426(20):3342–9.
10. Meyer-Ficca ML, Lonchar JD, Ihara M, Meistrich ML, Austin CA, Meyer RG. Poly(ADP-ribose) polymerases PARP1 and PARP2 modulate topoisomerase II beta (TOP2B) function during chromatin condensation in mouse spermiogenesis. Biol Reprod. 2011;84:900–9.
11. Gawecka JE, Ribas-Maynou J, Benet J, Ward WS. A model for the control of DNA integrity by the sperm nuclear matrix. Asian J Androl. 2015;17:610–5.
12. Gatewood JM, Cook GR, Balhorn R, Bradbury EM, Schmid CW. Sequence-specific packaging of DNA in human sperm chromatin. Science. 1987;236:962–4.
13. Jones EL, Zalensky AO, Zalenskaya IA. Protamine withdrawal from human sperm nuclei following heterologous ICSI into hamster oocytes. Protein Pept Lett. 2011;18:811–6.
14. Carrell DT, Hammoud SS. The human sperm epigenome and its potential role in embryonic development. Mol Hum Reprod. 2010;16:37–47.
15. Vavouri T, Lehner B. Chromatin organization in sperm may be the major functional consequence of base composition variation in the human genome. PLoS Genet. 2011;7(4):e1002036.
16. Rando OJ. Daddy issues: paternal effects on phenotype. Cell. 2012;151(4):702–8.
17. Jodar M, Oliva R. Protamine alterations in human spermatozoa. Adv Exp Med Biol. 2014;791:83–102.
18. Oliva R. Protamines and male infertility. Hum Reprod Update. 2006;12:417–35.
19. Castillo J, Amaral A, Azpiazu R, Vavouri T, Estanyol JM, Ballescà JL, et al. Genomic and proteomic dissection and characterization of the human sperm chromatin. Mol Hum Reprod. 2014;20:1041–53.
20. Balhorn R, Corzett M, Mazrimas JA. Formation of intraprotamine disulfides in vitro. Arch Biochem Biophys. 1992;296:384–93.
21. Björndahl L, Kvist U. Human sperm chromatin stabilization: a proposed model including zinc bridges. Mol Hum Reprod. 2010;16:23–9.
22. Balhorn R, Corzett M, Mazrimas J, Stanker LH, Wyrobek A. High-performance liquid chromatographic separation and partial characterization of human protamines 1, 2, and 3. Biotechnol Appl Biochem. 1987;9:82–8.
23. Nelson JE, Krawetz SA. Mapping the clonally unstable recombinogenic PRM1–>PRM2–>TNP2 region of human 16p13.2. DNA Seq. 1995;5:163–8.
24. Chirat F, Arkhis A, Martinage A, Jaquinod M, Chevaillier P, Sautière P. Phosphorylation of human sperm protamines HP1 and HP2: identification of phosphorylation sites. Biochim Biophys Acta. 1993;1203:109–14.
25. Castillo J, Estanyol JM, Ballescá JL, Oliva R. Human sperm chromatin epigenetic potential: genomics, proteomics, and male infertility. Asian J Androl. 2015;17:601–9.

26. Brunner AM, Nanni P, Mansuy IM. Epigenetic marking of sperm by post-translational modification of histones and protamines. Epigenetics Chromatin. 2014;7:1–12.
27. Pruslin FH, Imesch E, Winston R, Rodman TC. Phosphorylation state of protamines 1 and 2 in human spermatids and spermatozoa. Gamete Res. 1987;18:179–90.
28. Papoutsopoulou S, Nikolakaki E, Chalepakis G, Kruft V, Chevaillier P, Giannakouros T. SR protein-specific kinase 1 is highly expressed in testis and phosphorylates protamine 1. Nucleic Acids Res. 1999;27:2972–80.
29. Pirhonen A. Identification of phosphoseryl residues in protamines from mature mammalian spermatozoa. Biol Reprod. 1994;50:981–6.
30. de Mateo S, Ramos L, de Boer P, Meistrich M, Oliva R. Protamine 2 precursors and processing. Protein Pept Lett. 2011;18:778–85.
31. Luense LJ, Wang X, Schon SB, Weller AH, Lin Shiao E, Bryant JM, et al. Comprehensive analysis of histone post-translational modifications in mouse and human male germ cells. Epigenetics Chromatin. 2016;9:24.
32. Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, et al. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol. 2010;17:679–87.
33. Carrell DT, Hammoud SS. The human sperm epigenome and its potential role in embryonic development. Mol Hum Reprod. 2009;16:37–47.
34. van der Heijden GW, Ramos L, Baart EB, van den Berg IM, Derijck AA, van der Vlag J, et al. Sperm-derived histones contribute to zygotic chromatin in humans. BMC Dev Biol. 2008;8:34.
35. Nanassy L, Liu L, Griffin JT, Carrell D. The clinical utility of the protamine 1/protamine 2 ratio in sperm. Protein Pept Lett. 2011;18:772–7.
36. Corzett M, Mazrimas J, Balhorn R. Protamine 1: protamine 2 stoichiometry in the sperm of eutherian mammals. Mol Reprod Dev. 2002;61:519–27.
37. Balhorn R, Reed S, Tanphaichitr N. Aberrant protamine 1/protamine 2 ratios in sperm of infertile human males. Experientia. 1988;44:52–5.
38. Simon L, Liu L, Murphy K, Ge S, Hotaling J, Aston KI, et al. Comparative analysis of three sperm DNA damage assays and sperm nuclear protein content in couples undergoing assisted reproduction treatment. Hum Reprod. 2014;29:904–17.
39. Hamad MF, Shelko N, Kartarius S, Montenarh M, Hammadeh ME. Impact of cigarette smoking on histone (H2B) to protamine ratio in human spermatozoa and its relation to sperm parameters. Andrology. 2014;2:666–77.
40. Aoki VW, Liu L, Jones KP, Hatasaka HH, Gibson M, Peterson CM, et al. Sperm protamine 1/protamine 2 ratios are related to in vitro fertilization pregnancy rates and predictive of fertilization ability. Fertil Steril. 2006;86:1408–15.
41. Torregrosa N, Domínguez-Fandos D, Camejo MI, Shirley CR, Meistrich ML, Ballescà JL, et al. Protamine 2 precursors, protamine 1/protamine 2 ratio, DNA integrity and other sperm parameters in infertile patients. Hum Reprod. 2006;21:2084–9.
42. Mengual L, Ballescà JL, Ascaso C, Oliva R. Marked differences in protamine content and P1/P2 ratios. J Androl. 2003;24:438–47.
43. Khara KK, Vlad M, Griffiths M, Kennedy CR. Human protamines and male infertility. J Assist Reprod Genet. 1997;14:282–90.
44. de Yebra L, Ballescà JL, Vanrell JA, Bassas L, Oliva R. Complete selective absence of protamine P2 in humans. J Biol Chem. 1993;268:10553–7.
45. de Mateo S, Gázquez C, Guimerà M, Balasch J, Meistrich ML, Ballescà JL, et al. Protamine 2 precursors (pre-P2), protamine 1 to protamine 2 ratio (P1/P2), and assisted reproduction outcome. Fertil Steril. 2009;91:715–22.
46. Carrell DT, Liu L. Altered protamine 2 expression is uncommon in donors of known fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. J Androl. 2001;22:604–10.
47. Bach O, Glander HJ, Scholz G, Schwarz J. Electrophoretic patterns of spermatozoal nucleoproteins (NP) in fertile men and infertility patients and comparison with NP of somatic cells. Andrologia. 1990;22(3):217–24.

48. Lescoat D, Colleu D, Boujard D, Le Lannou D. Electrophoretic characteristics of nuclear proteins from human spermatozoa. Arch Androl. 1988;20:35–40.
49. Ribas-Maynou J, García-Peiró A, Martínez-Heredia J, Fernández-Encinas A, Abad C, Amengual MJ, et al. Nuclear degraded sperm subpopulation is affected by poor chromatin compaction and nuclease activity. Andrologia. 2015;47:286–94.
50. García-Peiró A, Martínez-Heredia J, Oliver-Bonet M, Abad C, Amengual MJ, Navarro J, et al. Protamine 1 to protamine 2 ratio correlates with dynamic aspects of DNA fragmentation in human sperm. Fertil Steril. 2011;95:105–9.
51. Castillo J, Simon L, de Mateo S, Lewis S, Oliva R. Protamine/DNA ratios and DNA damage in native and density gradient centrifuged sperm from infertile patients. J Androl. 2011;32:324–32.
52. Simon L, Castillo J, Oliva R, Lewis SEM. Relationships between human sperm protamines, DNA damage and assisted reproduction outcomes. Reprod BioMed Online. 2011;23:724–34.
53. Aoki VW, Emery BR, Liu L, Carrell DT. Protamine levels vary between individual sperm cells of infertile human males and correlate with viability and DNA integrity. J Androl. 2006;27:890–8.
54. Hammadeh ME, Hamad MF, Montenarh M, Fischer-Hammadeh C. Protamine contents and P1/P2 ratio in human spermatozoa from smokers and non-smokers. Hum Reprod. 2010;25:2708–20.
55. Nasr-Esfahani MH, Salehi M, Razavi S, Mardani M, Bahramian H, Steger K, et al. Effect of protamine-2 deficiency on ICSI outcome. Reprod BioMed Online. 2004;9:652–8.
56. Jodar M, Oriola J, Mestre G, Castillo J, Giwercman A, Vidal-Taboada JM, et al. Polymorphisms, haplotypes and mutations in the protamine 1 and 2 genes. Int J Androl. 2011;34:470–85.
57. Ni K, Spiess A-N, Schuppe H-C, Steger K. The impact of sperm protamine deficiency and sperm DNA damage on human male fertility: a systematic review and meta-analysis. Andrology. 2016;4:789–99.
58. Aoki VW. DNA integrity is compromised in protamine-deficient human sperm. J Androl. 2005;26:741–8.
59. Aoki VW, Liu L, Carrell DT. Identification and evaluation of a novel sperm protamine abnormality in a population of infertile males. Hum Reprod. 2005;20:1298–306.
60. Kouzarides T. Chromatin modifications and their function. Cell. 2007;128(4):693–705.
61. Ausió J. The shades of gray of the chromatin fiber. BioEssays. 2015;37:46–51.
62. Govin J, Caron C, Lestrat C, Rousseaux S, Khochbin S. The role of histones in chromatin remodelling during mammalian spermiogenesis. Eur J Biochem. 2004;271:3459–69.
63. Rajender S, Avery K, Agarwal A. Epigenetics, spermatogenesis and male infertility. Mutat Res. 2011;727:62–71.
64. Lin Q, Sirotkin A, Skoultchi AI. Normal spermatogenesis in mice lacking the testis-specific linker histone H1t. Mol Cell Biol. 2000;20:2122–8.
65. Fantz DA, Hatfield WR, Horvath G, Kistler MK, Kistler WS. Mice with a targeted disruption of the H1t gene are fertile and undergo normal changes in structural chromosomal proteins during spermiogenesis. Biol Reprod. 2001;64:425–31.
66. Martianov I, Brancorsini S, Catena R, Gansmuller A, Kotaja N, Parvinen M, et al. Polar nuclear localization of H1T2, a histone H1 variant, required for spermatid elongation and DNA condensation during spermiogenesis. Proc Natl Acad Sci. 2005;102:2808–13.
67. Tanaka H, Iguchi N, Isotani A, Kitamura K, Toyama Y, Matsuoka Y, et al. HANP1/H1T2, a novel histone H1-like protein involved in nuclear formation and sperm fertility. Mol Cell Biol. 2005;25:7107–19.
68. Fan Y, Sirotkin A, Russell RG, Ayala J, Skoultchi AI. Individual somatic H1 subtypes are dispensable for mouse development even in mice lacking the H10 replacement subtype. Mol Cell Biol. 2001;21:7933–43.
69. Changolkar LN, Costanzi C, Leu NA, Chen D, McLaughlin KJ, Pehrson JR. Developmental changes in histone macroH2A1-mediated gene regulation. Mol Cell Biol. 2007;27:2758–64.
70. Boulard M, Storck S, Cong R, Pinto R, Delage H, Bouvet P, et al. Histone variant macroH2A1 deletion in mice causes female-specific steatosis. Epigenetics Chromatin. 2010;3:8.

71. Celeste A, Petersen S, Romanienko PJ, Fernandez-Capetillo O, Chen HT, Sedelnikova OA, et al. Genomic instability in mice lacking histone H2AX. Science. 2002;296:922–7.
72. Faast R, Thonglairoam V, Schulz TC, Beall J, Wells JR, Taylor H, et al. Histone variant H2A.Z is required for early mammalian development. Curr Biol. 2001;11:1183–7.
73. Santoro SW, Dulac C, Banaszynski L, Allis C, Lewis P, Barski A, et al. The activity-dependent histone variant H2BE modulates the life span of olfactory neurons. Elife. 2012;1:662–74.
74. Tang MCW, Jacobs SA, Wong LH, Mann JR. Conditional allelic replacement applied to genes encoding the histone variant H3.3 in the mouse. Genesis. 2013;51:142–6.
75. Howman EV, Fowler KJ, Newson AJ, Redward S, MacDonald AC, Kalitsis P, et al. Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. Proc Natl Acad Sci. 2000;97:1148–53.
76. Kalitsis P, Fowler KJ, Earle E, Griffiths B, Howman E, Newson AJ, et al. Partially functional Cenpa–GFP fusion protein causes increased chromosome missegregation and apoptosis during mouse embryogenesis. Chromosom Res. 2003;11:345–57.
77. Maze I, Noh K-M, Soshnev AA, Allis CD. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. Nat Rev Genet. 2014;15:259–71.
78. Banaszynski LA, Allis CD, Lewis PW. Histone variants in metazoan development. Dev Cell. 2010;19:662–74.
79. Peterson CL, Laniel M-A. Histones and histone modifications. Curr Biol. 2004;14(14): R546–51.
80. Oliva R, Mezquita C. Histone H4 hyperacetylation and rapid turnover of its acetyl groups in transcriptionally inactive rooster testis spermatids. Nucleic Acids Res. 1982;10:8049–59.
81. Sonnack V, Failing K, Bergmann M, Steger K. Expression of hyperacetylated histone H4 during normal and impaired human spermatogenesis. Andrologia. 2002;34:384–90.
82. Govin J, Escoffier E, Rousseaux S, Kuhn L, Ferro M, Thévenon J, et al. Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. J Cell Biol. 2007;176:283–94.
83. Oliva R, Bazett-Jones D, Mezquita C, Dixon GH. Factors affecting nucleosome disassembly by protamines in vitro. Histone hyperacetylation and chromatin structure, time dependence, and the size of the sperm nuclear proteins. J Biol Chem. 1987;262:17016–25.
84. Hazzouri M, Pivot-Pajot C, Faure AK, Usson Y, Pelletier R, Sèle B, et al. Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. Eur J Cell Biol. 2000;79:950–60.
85. Faure AK, Pivot-Pajot C, Kerjean A, Hazzouri M, Pelletier R, Péoc'h M, et al. Misregulation of histone acetylation in Sertoli cell-only syndrome and testicular cancer. Mol Hum Reprod. 2003;9:757–63.
86. Godmann M, Auger V, Ferraroni-Aguiar V, Di Sauro A, Sette C, Behr R, et al. Dynamic regulation of histone H3 methylation at lysine 4 in mammalian spermatogenesis. Biol Reprod. 2007;77(5):754–64.
87. Glaser S, Lubitz S, Loveland KL, Ohbo K, Robb L, Schwenk F, et al. The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. Epigenetics Chromatin. 2009;2:5.
88. Payne C, Braun RE. Histone lysine trimethylation exhibits a distinct perinuclear distribution in Plzf-expressing spermatogonia. Dev Biol. 2006;293:461–72.
89. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. Distinctive chromatin in human sperm packages genes for embryo development. Nature. 2009;460:473–8.
90. Krejčí J, Stixová L, Pagáčová E, Legartová S, Kozubek S, Lochmanová G, et al. Post-translational modifications of histones in human sperm. J Cell Biochem. 2015;116:2195–209.
91. Siklenka K, Erkek S, Godmann M, Lambrot R, McGraw S, Lafleur C, et al. Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. Science. 2015;350(6261):aab2006.
92. Arpanahi A, Brinkworth M, Iles D, Krawetz SA, Paradowska A, Platts AE, et al. Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. Genome Res. 2009;19:1338–49.

93. Samans B, Yang Y, Krebs S, Sarode GV, Blum H, Reichenbach M, et al. Uniformity of nucleosome preservation pattern in mammalian sperm and its connection to repetitive DNA elements. Dev Cell. 2014;30:23–35.
94. Royo H, Stadler MB, Peters AHFM, Arpanahi A, Brinkworth M, Iles D, et al. Alternative computational analysis shows no evidence for nucleosome enrichment at repetitive sequences in mammalian spermatozoa. Dev Cell Elsevier. 2016;37:98–104.
95. Zhong HZ, Lv FT, Deng XL, Hu Y, Xie DN, Lin B, et al. Evaluating γh2AX in spermatozoa from male infertility patients. Fertil Steril. 2015;104:574–81.
96. Zhang X, Gabriel MS, Zini A. Sperm nuclear histone to protamine ratio in fertile and infertile men: evidence of heterogeneous subpopulations of spermatozoa in the ejaculate. J Androl. 2006;27:414–20.
97. Zini A, Gabriel MS, Zhang X. The histone to protamine ratio in human spermatozoa: comparative study of whole and processed semen. Fertil Steril. 2007;87:217–9.
98. Zini A, Zhang X, Gabriel MS. Sperm nuclear histone H2B: correlation with sperm DNA denaturation and DNA stainability. Asian J Androl. 2008;10:865–71.
99. Singleton S, Zalensky A, Doncel GF, Morshedi M, Zalenskaya IA. Testis/sperm-specific histone 2B in the sperm of donors and subfertile patients: variability and relation to chromatin packaging. Hum Reprod. 2007;22:743–50.
100. Hammoud SS, Nix DA, Hammoud AO, Gibson M, Cairns BR, Carrell DT. Genome-wide analysis identifies changes in histone retention and epigenetic modifications at developmental and imprinted gene loci in the sperm of infertile men. Hum Reprod. 2011;26:2558–69.
101. Azpiazu R, Amaral A, Castillo J, Estanyol JM, Guimerà M, Ballescà JL, et al. High-throughput sperm differential proteomics suggests that epigenetic alterations contribute to failed assisted reproduction. Hum Reprod. 2014;29:1225–37.
102. de Mateo S, Castillo J, Estanyol JM, Ballescà JL, Oliva R. Proteomic characterization of the human sperm nucleus. Proteomics. 2011;11:2714–26.
103. Baker MA, Naumovski N, Hetherington L, Weinberg A, Velkov T, Aitken RJ. Head and flagella subcompartmental proteomic analysis of human spermatozoa. Proteomics. 2013;13:61–74.
104. Castillo J, Amaral A, Oliva R. Sperm nuclear proteome and its epigenetic potential. Andrology. 2014;2:326–38.

7.2. Annex 2

La proteòmica del semen i la infertilitat masculina

Journal of Proteomics

2017 Jun 6;162:125-134

PMID: 27576136

Semen proteomics and male infertility
Jodar M, Soler-Ventura A, Oliva R.

Journal of Proteomics 162 (2017) 125–134

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Semen proteomics and male infertility

Meritxell Jodar *, Ada Soler-Ventura, Rafael Oliva *
Molecular Biology of Reproduction and Development Research Group

Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Faculty of Medicine, University of Barcelona, Casanova 143, 08036 Barcelona, Spain
Biochemistry and Molecular Genetics Service, Hospital Clínic, Villarroel 170, 08036 Barcelona, Spain

ARTICLE INFO

Article history:
Received 2 May 2016
Received in revised form 8 July 2016
Accepted 25 August 2016
Available online 26 August 2016

Keywords:
Proteomics
Sperm
Seminal fluid
Infertility
Seminal parameters
Epigenetics

ABSTRACT

Semen is a complex body fluid containing an admixture of spermatozoa suspended in secretions from the testes and epididymis which are mixed at the time of ejaculation with secretions from other accessory sex glands such as the prostate and seminal vesicles. High-throughput technologies have revealed that, contrary to the idea that sperm cells are simply a silent delivery vehicle of the male genome to the oocyte, the sperm cells in fact provide both a specific epigenetically marked DNA together with a complex population of proteins and RNAs crucial for embryogenesis. Similarly, –omic technologies have also enlightened that seminal fluid seems to play a much greater role than simply being a medium to carry the spermatozoa through the female reproductive tract. In the present review, we briefly overview the sperm cell biology, consider the key issues in sperm and seminal fluid sample preparation for high-throughput proteomic studies, describe the current state of the sperm and seminal fluid proteomes generated by high-throughput proteomic technologies and provide new insights into the potential communication between sperm and seminal fluid. In addition, comparative proteomic studies open a window to explore the potential pathogenic mechanisms of infertility and the discovery of potential biomarkers with clinical significance.

Significance: The review updates the numerous proteomics studies performed on semen, including spermatozoa and seminal fluid. In addition, an integrative analysis of the testes, sperm and seminal fluid proteomes is also included providing insights into the molecular mechanisms that regulate the generation, maturation and transit of spermatozoa. Furthermore, the compilation of several differential proteomic studies focused on male infertility reveals potential pathways disturbed in specific subtypes of male infertility and points out towards future research directions in the field.

Contents

* Corresponding authors.
E-mail addresses: m.jodar.bifet@gmail.com (M. Jodar), roliva@ub.edu (R. Oliva).

http://dx.doi.org/10.1016/j.jprot.2016.08.018
1874-3919/

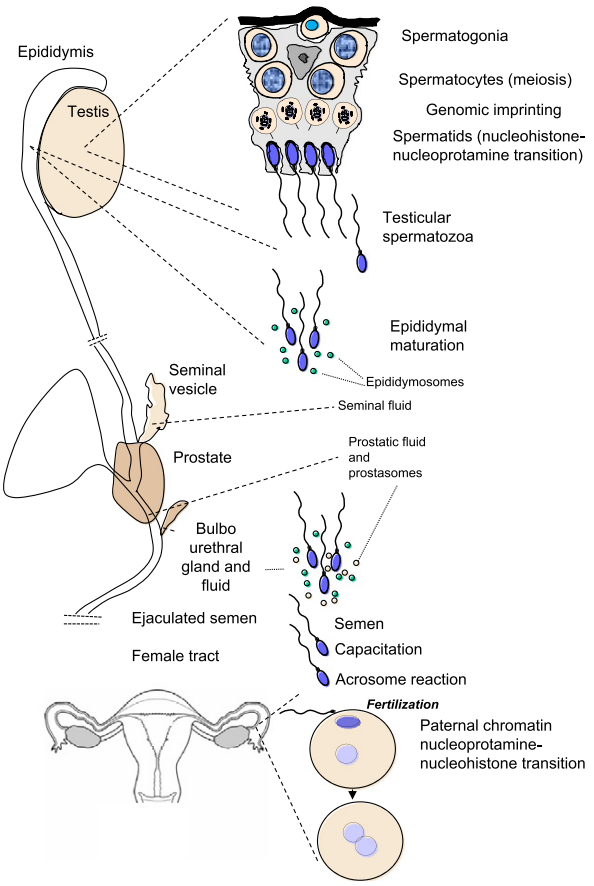


Fig. 1. Overview of the natural history of sperm generation, maturation and functional activities. Spermatogonia replicate and differentiate into spermatocytes, where meiotic recombination takes place. After meiosis a differentiation process called spermiogenesis takes place in which the round spermatid experiences profound cellular, chromatin and biochemical changes resulting in the generation of the highly differentiated sperm cell in the testes. Testicular spermatozoa initiate then a maturation process through it transit in the epididymis where they acquire the capacity to swim and additional membrane and biochemical functionalities. During ejaculation the sperm cells are further mixed with the seminal vesicles and prostatic fluids containing extracellular vesicles and with the bulbourethral gland fluid. Once in the female tract the sperm cells are capacitated acquiring hypermotility. Finally, acrosome reaction and fertilization takes place where the male paternal genome with its imprinting through DNA methylation, chromatin poising, RNAs and sperm protein cargo is transferred to the oocyte.

1. Introduction

Infertility is a growing problem that affects approximately 15% of the couples at reproductive age [1]. The causes of infertility are almost equally divided between male and female factors, each accounting for approximately 40% of all cases of infertility, with the remaining 20% is due to a combination of factors from both partners. The diagnosis of male infertility is currently based on the study of sperm quality including the analysis of seminal parameters such as sperm concentration, motility and morphology. To date, a limited number of causes have been associated with altered seminal parameters. This is exemplified by the dramatic decrease of sperm count caused for the presence of chromosome aberrations or microdeletions at the Y chromosome [2]. However, in the majority of male patients the molecular defect responsible for the infertility, remains unknown [3]. The origin may be due to the presence of genetic, epigenetic, systemic or environmental changes, or a combination of those [3–7]. High-throughput techniques, including genomics, epigenomics, transcriptomics, proteomics and metabolomics hold the promise to provide insights into the basic molecular mechanisms of sperm production and maturation as well as into the pathogenesis of male infertility [8,9]. Among all these levels, proteins are the main effectors of the function in cells. Therefore, the present review focuses on the recent application of proteomics to uncover the fundamental questions of the composition and function of the human sperm cell and on the discovery of altered molecular proteins and pathways in the seminal fluid and sperm cells of infertile patients. In addition, we have compiled, analyzed and discussed the integrative aspects of the currently available human testes, spermatozoa and seminal fluid proteomes.

2. Sperm cell biology

Spermatozoa are the culmination of the complex hormonally controlled process of spermatogenesis which takes place in the seminiferous tubules of the testes (Fig. 1) [10]. Spermatogenesis comprises three distinct phases: (i) During the first phase of mitotic proliferation, the undifferentiated diploid male germ cell differentiates into two different types of spermatogonia. Spermatogonia type A replicate by mitosis to ensure the maintenance of the future spermatogonial population while spermatogonia type B enter into meiosis and differentiate towards spermatozoa. (ii) During the 2nd phase, cytoplasm in type B cells increases, turning into primary spermatocytes. The spermatocyte undergoes two consecutive meiotic divisions thus reducing by half its genetic material and producing round haploid spermatids. (iii) Through the final phase, referred to as spermiogenesis, the round spermatids undergo significant nuclear, morphological and cytoplasmic changes to become mature spermatozoa [11].

Nuclear changes include the general replacement of histones by more basic nuclear proteins, the protamines [12–16]. This replacement results in a very high condensation of the germ cell nucleus, effectively blocking the transcription process. Two different morphological structures are formed during spermiogenesis, the acrosome and the sperm tail. Golgi vesicles fuse into one large vesicle that flattens and then surrounds two-thirds of the nuclear perimeter forming the acrosome. The acrosome contains proteases and other hydrolytic enzymes required to penetrate the zona pellucida during fecundation. Simultaneously, the centrosome migrates towards the opposite pole of the nucleus from the acrosome. Moreover, from one of the two centrioles present in the centrosome starts the formation of the flagellum axoneme. At the same time, mitochondria from round spermatids migrate primarily to this region and locate themselves to surround the proximal part of flagellum axoneme and form the mid part of the spermatozoa. Finally, the excess of cytoplasm is removed in residual bodies. Most of the cytoplasm is phagocytized by Sertoli cells but a carryover cytoplasmic droplet remains attached to the neck of the testicular spermatozoa to be removed during sperm maturation [17].

After successive stages of differentiation in the testes, the resultant spermatozoon is still only an immature germ cell being as yet immotile and lacking the ability to fertilize the oocyte on its own. Sperm maturation, acquiring both motility and fertilization potential, progresses through its transit along the epididymis over the span of one or two weeks and to a further extent upon entry into the female tract (Fig. 1).

Testicular spermatozoa are stored in a coiled tube structure located in the back of the testes, the epididymis, as well as in the proximal portion of the vas deferens, near the epididymis, until ejaculation. The epithelial cells of epididymis secrete vesicles, specifically named epididymosomes, to the epididymis lumen which are rich in lipids, proteins and RNAs that will interact with the sperm membrane resulting in sperm surface changes and the acquisition of new proteins and RNAs [18]. These sperm changes contribute to the acquisition of the potential for sperm motility and the ability to fertilize the oocyte as well as probably to transfer epigenetic transgenerational information [19,20]. In addition to these maturation steps, the epididymis also provides a complex microenvironment which maintains the sperm cell in a quiescent state during its epididymal storage and avoid a premature sperm activation [21].

During the ejaculation, sperm travels from the epididymis through the vas deferens and arrive first at the ampulla, where secretions from the seminal vesicles are added (~70% of seminal fluid volume). From the ampulla, the seminal fluid flows through the ejaculatory ducts towards the urethra, passes to the prostate gland where prostate secretions are added (~20% of seminal fluid volume) and finally semen is secreted outside the body. Seminal vesicles and prostate secretions contribute with: (i) sugars essential for spermatozoa nutrition during its transit to the oocyte [22], (ii) proteins required for the coagulation and liquefaction of semen such as semenogelins I and II and kallikrein 3 proteins secreted from the seminal vesicles and prostate, respectively, and (iii) immune components. Moreover, epithelial cells from prostate gland also secrete a high number of vesicles named prostasomes with a cargo of different lipids, proteins, RNA and DNA. There are plenty of evidences showing the role of prostasomes in sperm function as for example the stimulation of the sperm motility [23], the regulation of the capacitation timing and also the sperm protection against the immune response in the female tract [24]. Bulbourethral and urethral glands secrete several mucoproteins representing a small percentage of the fluid. These classes of proteins typically have roles in intercellular interaction and determination of immune properties.

The last steps of sperm maturation occur within the female tract. After ejaculation, the spermatozoa which survive the cervical mucus, begin their ascension up to the oviduct. Exposure to different female secretions within the vagina activates the sperm to initiate the process of capacitation. During capacitation there are general changes in the

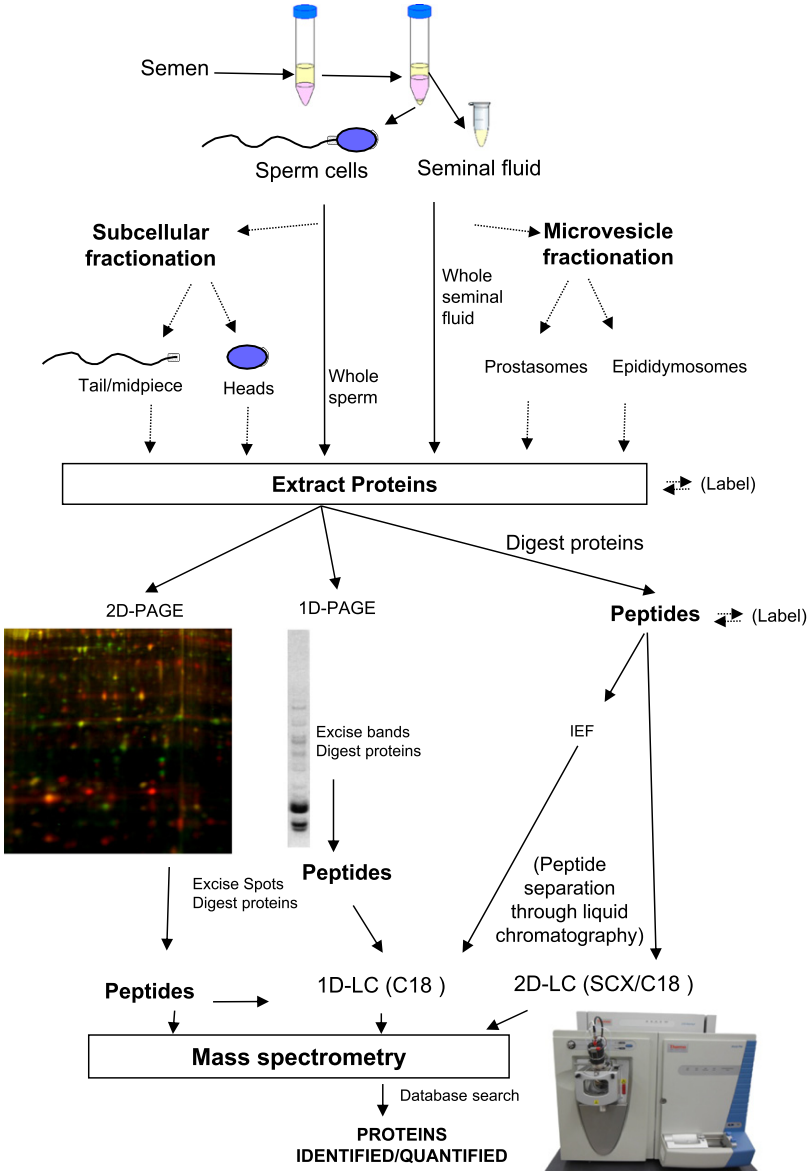


Fig. 2. Basic methodological steps in the proteomic analysis of the sperm cells and other seminal components. Ejaculated semen is a mixture composed by a majority of sperm cells (5×10^6–10^8/ml), a small proportion of somatic cells (leukocytes, epithelial cells), and the seminal fluid components including prostasomes and epididymosomes. So depending on the experimental goal it is essential to purify first the desired components. This is usually accomplished through colloidal sillica density gradients centrifugation or through swim-up. Also subcellular fractionation or seminal fluid fractionation is possible as alternatives. Once the desired target components have been purified the next step is to extract the proteins. Proteins can then be separated by gel electrophoresis (left and center) and the desired proteins eluted and digested into peptides. Alternatively, the original protein mixture can be directly digested into peptides. The final stage is to separate the peptides through liquid chromatography and to proceed to its identification using mass spectrometry.

sperm surface including the loss of specific membrane proteins and the modification of the cholesterol:phospholipid ratio [25]. All these changes serve to increase the permeabilization of the sperm surface promoting the hyperactive motility of spermatozoa and their ability to undergo the acrosome reaction upon reaching the oocyte. The interaction of sperm surface proteins complementary to zona pellucida proteins enables the acrosomal reaction. The membrane surrounding the acrosome fuses with the plasma membrane of the spermatozoa releasing the enzyme content responsible for the digestion of the oocyte membrane. The membrane of the fertilizing spermatozoon fuses with the oocyte membrane and the chromatin, RNA, and protein contents of the sperm head are transmitted into the egg.

Proteomic studies on sperm and seminal fluid from different types of patients could help to clarify some of the molecular mechanisms involved in the generation of fertile spermatozoa. For example, the semen proteomics studies on individuals with spinal cord injury shed new light on prostate gland functions [23].

3. Sample preparation for semen proteomic studies

Semen is a complex fluid comprising a cellular and a non-cellular fraction. The cellular components are mainly the spermatozoa which accounts for approximately 5% of the semen volume, whereas the acellular fraction or seminal fluid contains secretions from different accessory sex glands which account for 95% of semen. Accessory sex glands not only secrete fluid to semen, but also release trillions of extracellular vesicles (10^{11} to 10^{13} particles) rich in DNA, RNA, lipids and proteins being the prostate the major contributor [26]. Therefore, proteomics studies can be performed on sperm, seminal fluid and seminal extracellular vesicles.

For sperm proteomics studies, the spermatozoa should be purified from a minor proportion of somatic cell types such as leukocytes and epithelial cells and immature germ cells usually present in semen (Fig. 2). The common purification method used routinely in the clinics is the sperm purification through centrifugation with colloidal silica density gradients. Alternatively, depending on the desired goal, additional methods are available such as swim-up to isolate the motile sperm cells, CD45-magnetic beads for leukocyte depletion, among other alternatives [27–29].

In contrast, in order to obtain the seminal fluid, the sperm and other potentially contaminating cells should be removed by simple centrifugation. The extracellular vesicles present in human seminal fluid could be collected from ejaculates removing the spermatozoa by simple centrifugation followed by different ultracentrifugation steps of the supernatant [30]. Most of these extracellular vesicles are secreted from the prostate, for this reason, some studies that use this methodology, assumes the isolation of a pure population of prostasomes. However, the seminal fluid ultracentrifugation does not rule out the presence of extracellular vesicles released by other glands.

Though it is known that the epididymis releases a high number of extracellular vesicles specifically named epididymosomes which interact with sperm membrane during epididymal storage, some studies suggest that the epididymosomes do not contribute significantly to the population of vesicles present in the ejaculate, at least in bovines [31]. In order to isolate human epididymosomes, they should be collected from the open end of the scrotal portion of the vas deferens while a vasectomy reversal surgery is done. In order to purify the epididymosomes, contaminating cells and cellular debris should be removed from the epididymal fluid by regular centrifugation. After this step, the supernatant is then ultracentrifuged twice and the epididymosomes are collected [32].

Once the biological material has been appropriately prepared (sperm, seminal fluid or extracellular vesicles) the next step is to proceed to the extraction of the proteins and to their identification [33]. The separation of proteins or peptides is essential when performing large scale assays of cell extracts or complex samples. This can be accomplished with SDS-PAGE separation, in which proteins are separate by their molecular weight in gels with ionic detergents such as SDS or with proteins isoelectrofocusing. Both methods for protein separation have a low resolution if applied independently. However, when combined in the 2D-PAGE approach the resolution of separation increases considerably. This method separates thousands of proteins as spots in a 2D gel which can be isolated, in-gel digested and analyzed by mass spectrometry.

These 2D-PAGE based approaches have been extensively used for quantitative analysis in sperm studies but presently are being replaced by higher throughput approaches. This approaches relies on the separation of peptides (rather than proteins) using liquid chromatography (LC) followed by peptide identification using tandem mass spectrometry (MS/MS) [34]. In a LC-MS/MS experiment, the peptides generated after a specific digestion of the proteins by proteases (such as trypsin), are usually separated based on their hydrophobicity by using reverse phase columns (RPLC) and, once eluted, they are identified using MS/MS. More recently, additional chromatographic approaches have been coupled prior to the LC-MS/MS such as ionic exchange chromatography, which separates peptides by their charge prior to the RPLC and MS/MS. The combination of methodologies based on different chemical properties is referred as two-dimensional liquid chromatography coupled to MS/MS (2D-LC-MS/MS) [35].

4. Semen proteomics

Spermatozoa are a very suitable model for proteomic analysis since they can be easily purified from semen. Furthermore, spermatozoa do not generate new proteins encoded by nuclear genes, as they are transcriptionally and translationally inert cells due to their high DNA compaction and ejection of the translation machinery during the late stages of spermatogenesis. That fact, reduces the complexity of the sperm proteome analysis [36]. In contrast, the composition of the seminal fluid is very heterogeneous as it is formed by a mixture of proteins released from different organs resulting in a cell-free fluid or in an encapsulated fluid as extracellular vesicles. The rapid changes that occur in the seminal fluid after ejaculation, such as the coagulation-liquefaction process, introduce further complexity to seminal fluid proteomic studies. These caveats are reflected by the higher number of studies assessing the spermatozoon proteome as compared to the seminal fluid proteome, as it will be discussed below.

4.1. Spermatozoal proteome

Proteomic spermatozoa profiles for this review were acquired from the comprehensive review of Amaral *et al.* (2014) [28] that included thirty different proteomic studies performed on human sperm which were available online until the end of 2012. The study covered the whole sperm proteomics, subcellular proteomics, differential proteomics involving sperm cells from infertile patients and functional proteomics such as proteins involved in the capacitation process [4]. Additionally, another proteomic characterization of head and tail sperm subcellular fractions was included [37]. Sperm nuclei proteome included the characterized sperm nuclei [38], the head subcellular fraction [37], and the characterized soluble and insoluble chromatin proteome [27]. In order to describe the sperm tail proteome, the combined proteome of the fibrous sheath of the human sperm flagellar principal piece [39] and the tail subcellular fraction [28,37] were incorporated. Only proteins with experimental evidence of their existence were considered in the current review.

All proteomic studies included in this review have resulted in the identification with high confidence of 6238 non-redundant proteins in the entire spermatozoa. Whereas the majority of the proteins (3824) were not identified in any of the proteomic studies from isolated sperm subcellular fractions, approximately 40% of these proteins are exclusively located in head (898 proteins) or tail (984 proteins) or indistinctly in both subfractions (532 proteins) (Fig. 3A). Ontological

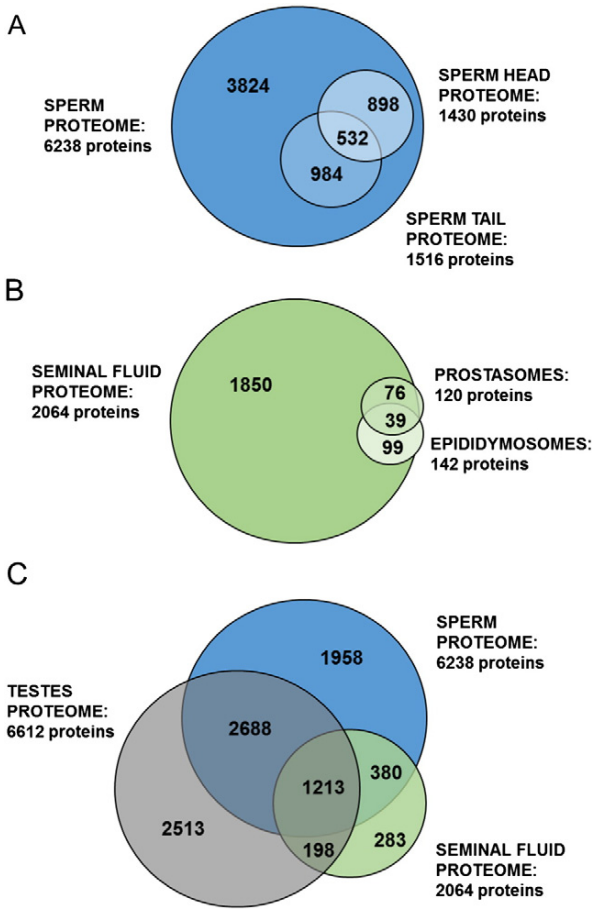


Fig. 3. Overview of the proteomes corresponding to the testes, sperm cells and seminal fluid. A: Currently available sperm cell proteome is formed by 6238 proteins and includes the sperm head and sperm tail proteomes. B: A total of 2064 proteins have been identified so far in the seminal fluid. Upon isolation and proteomic analysis of the particles present in the seminal fluid a total of 120 have been reported in the prostasomes and a total of 142 proteins have been identified in the epididymosomes. C: Venn diagram of the intersection of the testicular, sperm and seminal fluid proteomes.

analysis of GO component, biological process, molecular function and KEGG pathway were assessed using DAVID bioinformatics resources v6.7 (https://david.ncifcrf.gov/) [40,41] for each subset of proteins based on their subcellular localization (Supplementary Table 1). Only the GO term and KEGG pathways significantly overrepresented after the multiple test comparison correction (Bonferroni correction) were considered ($p < 0.05$).

As expected, sperm proteins exclusively present in sperm head are enriched in processes such as DNA packaging, RNA metabolism (splicing and transport) and nucleocytoplasmic transport [42]. Surprisingly, other processes known to be blocked in mature sperm such as DNA replication and DNA repair [43] are also highly enriched in sperm head, most probably as remnants of spermatogenesis. In contrast, the proteins found exclusively in the tail are enriched in processes required for the generation of precursor metabolites and energy. Specifically, the oxidative phosphorylation pathway (OXPHOS) that takes place in mitochondria is the most overrepresented pathway, although other processes to obtain energy such as citrate cycle (TCA cycle) and fatty acid metabolism have also been found overrepresented, as described previously in Amaral *et al.* (2014) [28]. Components from the cytoskeleton and more specifically axonemal components and the associated proteins required for the flagella beating, such as the dynein complex, are also specifically enriched in sperm tail.

A small number of proteins (532 proteins) are indistinctively detected in sperm head and tail. On the one hand, these proteins could represent a nonspecific accumulation of sperm head and tail proteins during the subfractionation process. This could be exemplified for proteins typically described as nucleosomal, mitochondrial or flagellar proteins. On the other hand, this common population may indicate a broad localization of the proteins along the sperm cell including its substructures. This is the case of the ubiquitous ubiquitin-proteasome complex with crucial roles in chromatin remodeling and biogenesis of sperm head and tail occurred during spermatogenesis [44], as well as, through its participation in sperm capacitation, acrosome reaction, egg-sperm interaction and processing of sperm proteins and organelles after fertilization [45,46] . Indirect immunofluorescence of specific proteins from ubiquitin-proteosome complex validate this observations since they were detected in the sperm head, particularly in the post-acrosomal region [47], and in the outer dense fibers of the sperm tail [48].

Proteins comprised in sperm ATP producing pathways - the glycolysis and the citrate cycle (TCA) - are also located indistinctively in both sperm head and tail fraction. Whereas it is known that glycolysis takes place in sperm head and along the principal piece of flagellum [49], the TCA cycle is commonly defined as a mitochondrial process. However, it would not be surprising that some of these sperm proteins detected "out of the expected location" turn out to be genuine functional proteins in those "ectopic" locations due to the extremely peculiar nature of the sperm cell.

In addition to the proteins with a known sublocalization, a high percentage of the sperm proteins were not detected in neither the head or the tail proteomes suggesting that they were lost during sperm head or tails isolation processes. The enrichment of proteins usually associated to the extracellular region corroborates this hypothesis (Supplementary Table 1). Surprisingly, cytosolic ribosomal proteins also seem to be lost during the subfractionation processes, most probably due to their association to membranes or vesicles.

4.2. Seminal fluid proteome

A review containing nine different studies was used as a source of the seminal fluid proteome [50]. Only one study for each type of seminal fluid extracellular vesicles (epididymosomes and prostasomes) was detected after a systematic search with the following keywords "prostasome proteomics"; "epididymosomes proteomics"; "prostasome proteins"; "epididymosomes proteins" in the PubMed database [32,30].

A total of 2064 non-redundant proteins were identified in the seminal fluid, however this number of proteins is considerably lower than those identified in other body fluids, as for example the 10,000 proteins detected in blood plasma. Preliminary results suggest that the 10 more abundant proteins in seminal fluid including the semenogelins I and II (SEMG1 and SEMG2) represent the 80% of all proteins bulk mass [51]. The large dynamic range of the protein abundance in the seminal fluid may hinder the detection of the low abundance proteins. Therefore, strategies to deplete high abundant proteins, could improve the proteome characterization of seminal fluid, such as: (i) HPLC columns containing antibodies to the most abundant proteins or (ii) protein equalization methods to reduce the dynamic range of the proteins rather than to deplete them. Until now, only one study used an equalization method to assess the seminal fluid proteome indentifying 61 new proteins [52].

A small percentage (10%) of detected seminal fluid proteins are encapsulated in the seminal extracellular vesicles including those ones released by epididymis (epididymosomes) and prostate (prostasomes) (Fig. 3B). Enrichment analysis of GO terms summarized in Supplementary Table 2 revealed that proteins commonly enriched in prostasomes and epididymosomes are involved in molecular processes such as GTPase activity and calcium-dependent phospholipid binding which are known to be important for the release, trafficking and fusion of extracellular vesicles [53]. Beside

these basic pathways in the biogenesis of extracellular vesicles, other processes were also overrepresented in both types of extracellular vesicles as for example the glycolysis pathway. The known ability of the prostasomes to produce extracellular ATP from carbohydrates [54] suggest that seminal extracellular vesicles could act as an auxiliary instrument to provide energy to the sperm and thus promote sperm motility [55]. Prostasomes and epididymosomes are also enriched in antiapoptotic factors, although their potential role in apoptosis regulation was never assessed. Some proteins crucial for fertilization such as Cysteine-Rich Secretory Protein 1 (CRISP1) [56] were found exclusively in epididymosomes. These finding enhance the idea that epididymis provide some proteins to sperm required to fertilize the oocyte. In contrast, no unique enriched pathways were detected for prostasomes, maybe reflective of the limitations in the methodology used to isolate them. Whereas the epididymosomes were isolated from epididymal fluid, the prostasomes were purified from ejaculated samples by differential ultracentrifugation. As mentioned above, this methodology is not specific to isolate prostasomes since it does not rule out the presence of extracellular vesicles released by other glands, apoptotic bodies or other cell debris.

Additionally, a complex mixture of relevant biological pathways is observed for the seminal fluid proteins not encapsulated in the extracellular vesicles (Supplementary Table 2). Some of the enriched pathways and processes have already been described as specific from seminal fluid such as the overrepresentation of complement and coagulation pathways [6]. However, biological pathways usually enriched in spermatozoa were also enriched in seminal fluid, suggesting a nonspecific contamination. Most probably, the presence of sperm proteins in seminal fluid is reflective of sperm that undergone acrosome reaction or apoptosis. An integrative analysis of sperm, seminal fluid and testes proteome presented below might clarify the origin of these free proteins present in seminal fluid.

4.3. Integrative analysis of testes, spermatozoa and seminal fluid proteome

The integrative analysis of testes, spermatozoa and seminal fluid proteomes could be useful to provide new insights on the biology of the generation and maturation of sperm cells as well as on the potential communication pathways between sperm and its environment [6,57,20]. A database based on the testes proteins described from a unique healthy man was employed to describe the testes proteome [58]. As observed in Fig. 3C the proteome of testes, spermatozoa and seminal fluid was overlapped.

The junction of the three different proteomes enables to discern the sperm proteins probably originated in the testes during spermatogenesis (sperm proteins also present in the testes proteome; 3901 proteins) from those ones most likely imported from seminal fluid (Sperm proteins absent in the testes but detected in seminal fluid proteome; 380 proteins) (Fig. 3C).

The presence of proteins in both sperm and testes proteomes (3901 proteins) reflect their selective retention in spermatozoa during spermatogenesis, likely for their potential role in sperm function. This is exemplified by the enrichment of KEGG pathways related to the energy production (OXPHOS, fatty acid metabolism and citrate cycle) and axoneme and centrosome components required for sperm motility and embryo development [59], respectively. Approximately 30% of these sperm proteins (1213 proteins), most of them related to acrosome and proteasome complexes, are also present in the seminal fluid proteome suggesting that they could be remnants in seminal fluid from sperm that underwent the acrosome reaction or apoptosis. In contrast, proteins that are not essential for sperm function as for example proteins involved in blocked pathways in mature spermatozoa such as cell cycle, DNA replication, DNA repair, RNA transcription and RNA processing are exclusively detected in testes proteome (2513 proteins). An additional set of proteins appear as unique in the mature sperm (1958 proteins) suggesting that they could be either minor testicular proteins undetectable in the testes through high-throughput proteomic studies but selectively retained in sperm, or sperm proteins taken up from the epididymal fluid and acquired during the sperm maturation. Interestingly, the majority of the proteins exclusively detected in the sperm proteome are integral to the plasma membrane and are involved in calcium signaling pathways essential for sperm motility [60] or ABC transporters with a pivotal role in the cholesterol homeostasis during sperm maturation in epididymis [61].

After the exclusion of the 1213 seminal fluid proteins detected in all three proteomes (testes, sperm and seminal fluid) representing most probably remnants of activated sperm, the remaining 861 seminal fluid proteins could be divided in two different subgroups:

(i) 380 seminal fluid proteins also detected in the sperm proteome suggestive of an active transport to sperm. The proteins potentially incorporated to sperm from seminal fluid are enriched in molecular processes such as calcium ion binding, secretory granule, defense response to bacterium, cell motility and peptidase activity required for sperm motility and capacitation (Supplementary Fig. 1).
(ii) 481 seminal fluid proteins absent in spermatozoa. Although 40% of these proteins (198 proteins) were also detected in testes proteome, these are enriched in the broad distributed complement and coagulation pathways, likely participating in the semen clotting and liquefaction. Seminal fluid specific proteins are enriched in molecular processes such as peptidase inhibitor activity and glycosaminoglycan (GAG) binding proteins specifically for those ones binding heparin.

Semenogelin I and II released by seminal vesicles are the predominant proteins of the semen and vest the gel matrix appearance to semen coagulum after ejaculation. The proteolysis of these semegenolins results in the liquefaction of semen [62]. A high number of proteolysis regulators (proteases and their inhibitors), such as peptidases, are detected in sperm and seminal fluid proteome. Whereas the proteases are found in both sperm and seminal fluid, their inhibitors are only enriched in the extracellular fraction. Seminal fluid specific proteome also present an enrichment of glycosaminoglycan (GAGs) binding proteins, known to their ability to inhibit sperm capacitation [63]. Hence these observations suggest that the proteins solely detected in seminal fluid are crucial for maintenance of spermatozoa in a quiescent state during their storage.

The results from the integrative analysis are consistent with the view that all sperm functions required for fertilization should be poised in the mature spermatozoa but remain inactive until after ejaculation, where liquefaction increases the fluidity of the seminal fluid and allows the activation of all sperm functions necessary for successful transit and fertilization of the oocyte.

5. Discovering potential causes of male infertility

Approximately 40–60% of infertile males present some alteration in at least one of the seminal parameters (sperm concentration, motility and morphology) [64]. Oligozoospermia is characterized by a sperm concentration or total number of ejaculate sperm below 15 or 39 millions of sperm cells respectively, asthenozoospermia is considered when <32% of the spermatozoa have progressive motility and teratozoospermia is diagnosed when <4% of the sperm have normal morphology evaluated using Kruger's strict criteria [65]. Usually the alteration of the three seminal parameters appears combined in a single individual; e.g. oligozoospermia is often accompanied by poor motility (asthenozoospermia) and abnormal morphology (teratozoospermia). The comparative semen proteomic studies from infertile patients presenting

suboptimal semen parameters might indicate different pathways disrupted reflective of the altered semen parameters, therefore providing insight into the possible pathological mechanisms causing the different subtypes of male infertility.

A systematic search including the keywords "asthenozoospermia proteomics"; "teratozoospermia proteomics"; "oligozoospermia proteomics"; "oligoasthenoteratozoospermia proteomics"; "asthenozoospermia proteins"; "teratozoospermia proteins"; "oligozoospermia proteins"; "oligoasthenoteratozoospermia proteins" in Pubmed database resulted in a total of 25 studies assessing differential proteins or alterations in the proteomic profile of spermatozoa (22) or seminal fluid (3) from infertile patients presenting suboptimal seminal parameters (Supplementary Table 3). Briefly, the initial studies using proteomic technologies were based on several technical variations either in protein separation step including 2DE-SDS-PAGE and 2D-DIGE and in their corresponding identification (MALDI-TOF or LC-MS/MS) [66–77]. Likewise, the immobilized metal ion affinity chromatography (IMAC) method with a phosphoprotein enrichment kit followed by nano-UPLC-MS was used to describe differentially expressed sperm proteins in asthenozoospermic patients [78]. Proteomic analysis carried out by differential labeling followed by LC-MS/MS, allowed to increase the number of differentially expressed sperm proteins in asthenozoospermic patients [79]. Alternatively, another non-proteomic methodologies to detect differentially expressed sperm proteins were also considered [80–88] such as an enzyme-Linked immunosorbent assay (ELISA), immunofluorescence, enzymatic activity, flow cytometry, immunochemistry and western blot.

Differential proteomic studies revealed that semen including sperm and seminal fluid from oligozoospermic, asthenozoospermic and teratozoospermic patients present alterations of 16, 284 and 18 proteins, respectively (Supplementary Table 4). The higher number of proteins detected as differential in asthenozoospermia is reflective of the higher number of studies assessing this phenotype (18 studies for asthenozoospermia, 3 for oligozoospermia and 1 for teratozoospermia; Supplementary Table 3). Three additional studies evaluated combining phenotypes such as oligoasthenozoospermic and oligoasthenoteratozoospermic patients identify 11 differential proteins (Supplementary Table 4).

From the 284 differential proteins detected in asthenozoospermia, 193 correspond to sperm and 91 to seminal fluid proteins (Supplementary Table 4). As observed in Table 1, altered sperm proteins in asthenozoospermia mainly reflect a disturbance in the generation of precursor metabolites and energy, being the citrate cycle (TCA) the most affected pathway as previously observed by metabolomics studies [89]. According to previous findings showing a higher rate of apoptosis and necrosis in asthenozoospermia [90], differential proteomic studies in these type of patients also identify alterations in some inhibitory apoptotic proteins. Additionally, asthenozoospermic patients present alterations in specific proteins involved in sexual reproduction, specifically for those ones required for sperm motility as for example AKAP4, GAPDHS and ODF2 [91–93] (Supplementary Fig. 2). Similarly, the proteomic changes in the seminal fluid from asthenozoospermic patients are associated with the energy production, but being the glycolysis and not the TCA cycle the pathway affected [79]. Surprisingly, alterations in proteins usually related to the acrosome structure are also detected in patients with lower sperm motility (Table 1).

The low number of proteins detected as differential in oligozoospermic and teratozoospermic patients does not enable the identification of any disturbed pathway or process. Even so, it is interesting to note the alteration of two members from the SPANX gene family in teratozoospermic patients which have been potentially related to spermatogenesis [94].

6. Clinical perspectives of semen proteomics

Nowadays, routine assessment of male fertility in reproductive clinics is limited to examine semen parameters. Although assessment of the sperm concentration, motility and morphology resolves some severe male infertility factors, these are of little predictive value with respect to pregnancy outcome of different assisted reproductive technologies (ART). This observation emphasizes the need to develop molecular biomarkers able to discern the male factor infertility and in turn to be predictive of the success of different fertility treatments. So far, high throughput proteomics technologies only can be used as a biomarker discovery tool. However, once the clinical value for some of the candidate biomarkers detected by proteomics is validated, it should

Table 1
Pathways and processes affected in asthenozoospermic patients discovered by proteomic approaches. Differential sperm and seminal fluid proteins in asthenozoospermic patients revealed several Gene Ontologies processes and KEGG pathways significantly altered ($p < 0.05$ after Bonferroni correction).

	Number of genes	*p*-Value	Gene names	Bonferroni correction
Asthenozoospermic differential sperm proteins				
GO term				
GO:0006091—generation of precursor metabolites and energy	24	1,30E-12	ALDOA, DLST, NDUFB5, ECH1, ACO2, ATP5B, PGAM2, DLAT, PDHB, NDUFA11, SDHA, GAPDHS, SDHB, GPI, TPI1, GOT1, GSK3B, SDHC, DLD, SDHD, GAA, IDH1, ATP5O, FH	1,96E-09
GO:0006916—anti-apoptosis	14	7,47E-07	CDK1, YWHAZ, EEF1A2, CLU, PRDX2, ANXA5, HSP90B1, SH3GLB1, APOE, GSK3B, IL1B, HSPB1, HSPA5, ARHGDIA	1,13E-03
GO:0019953—sexual reproduction	20	1,35E-06	SEMG1, PTGS2, PGAM2, THEG, ELSPBP1, GAPDHS, KRT9, SPANXB1, CCT6B, HSPA2, DYNLL1, ROPN1B, GPX4, DLD, ODF2, KDM3A, ODF1, RUVBL1, AKAP3, AKAP4	2,04E-03
GO:0006986—response to unfolded protein	9	1,50E-06	HSPA1L, HSP90AB1, HSP90AA1, HSPA2, HSPA6, CSNK2B, HSPB1, DNAJB6, HSPA8	2,27E-03
KEGG pathways				
hsa00020: citrate cycle (TCA cycle)	11	1,27E-10	SDHA, DLST, SDHB, ACO2, SDHC, DLD, SDHD, IDH1, DLAT, PDHB, FH	1,44E-08
hsa00190: oxidative phosphorylation	10	7,55E-04	SDHA, SDHB, NDUFB5, ATP5B, SDHC, SDHD, ATP5O, COX6B2, COX5B, NDUFA11	8,18E-02
hsa00010: glycolysis/gluconeogenesis	7	8,99E-04	ALDOA, GPI, TPI1, DLD, PGAM2, DLAT, PDHB	9,67E-02
Asthenozoospermic differential seminal fluid proteins				
GO term				
GO:0006096—glycolysis	6	6,25E-06	ALDOA, GPI, LDHB, GAPDH, MDH1, ENO1	5,22E-03
GO:0006508—proteolysis	18	5,91E-05	CNDP2, LGMN, ANPEP, HP, NPEPPS, MYH9, QPCT, PRSS8, ACE, VCP, UBA1, PSMB2, KLK11, CTSB, ADAM7, CTSH, DPP4, WFDC2	4,83E-02
KEGG pathways				
hsa04142: lysosome	10	1,58E-07	NAGLU, GM2A, IGF2R, LGMN, SORT1, CTSB, CD63, CTSH, MANBA, GLB1	8,35E-06
hsa00010: glycolysis/gluconeogenesis	6	8,18E-05	ALDOA, GPI, LDHB, AKR1A1, GAPDH, ENO1	4,32E-03

be possible to develop cheaper and more feasible tests such as those based on protein microarrays, mass spectrometry selective reaction monitoring (SRM) or ELISA multiplexed to routinely test these specific biomarkers in the reproductive clinics.

So far, only three comparative semen proteomic studies have identified potential protein biomarkers indicative of pregnancy outcome after ART [95–98]. Further validation of these biomarkers could improve the reproductive counseling of infertile couples. Additionally, semen proteomics studies might also help to predict the success of some clinical procedures as for example: (i) successful sperm retrieval after testicular biopsy for azoospermic patients [99] or (ii) the improvement of semen quality after varicocelectomy procedure [100–102].

7. Concluding remarks

Extensive and accurate proteomic analysis have been performed in human spermatozoa resulting in the identification of a vast number of sperm proteins corresponding to the entire sperm cell proteome or to the different sperm substructures. However, until now only few studies are focused in the proteomic analysis of the seminal fluid. Integrative analysis of the testes, sperm and seminal fluid proteomes indicate the presence of a "communication" between sperm and seminal fluid but also a nonspecific mixture of the sperm and seminal fluid proteomes. These seminal fluid proteins most probably derive from the sperm cells that have undergone the acrosome reaction or apoptosis and may be interfering with the characterization of the proteins released by accessory sex glands. Further proteomic experiments such as MALDI imaging mass spectrometry from different epididymal segments or the proteomic characterization of the extracellular vesicles selected according to their origin may improve the present understanding of the impact of accessory sex glands on sperm maturation and male fertility. Moreover, integrative approaches comparing the proteins present in spermatozoa with the transcriptomic profiles of the organs involved in the production of semen may infer their origin and transcriptional timing [6,52].

Comparative proteomics have revealed specific biological pathways affected in infertile patients as for example the disturbance in the generation of energy mainly by the citrate cycle and also in the regulation of apoptosis in asthenozoospermic patients. The knowledge of the causes affecting the specific sperm parameters or the fertility could derive in the design of new therapies to improve male fertility as well as the discovery of molecular biomarkers with valuable clinical significance.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2016.08.018.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgements

This work was supported by grants from the Ministerio de Economía y Competitividad (FEDER, PI13/00699) from Fundación Salud 2000 (SERONO 13-015), EUGIN-UB (EU-REP 2014), and from EU-FP7-PEOPLE-2011-ITN289880 to RO.

References

[1] M.C. Inhorn, P. Patrizio, Infertility around the globe: new thinking on gender, reproductive technologies and global movements in the 21st century, Hum. Reprod. Update 21 (2014) 411–426, http://dx.doi.org/10.1093/humupd/dmv016.
[2] S. Pfeifer, S. Butts, D. Dumesic, G. Fossum, C. Gracia, A. La Barbera, R. Odem, M. Pisarska, R. Rebar, R. Reindollar, M. Rosen, J. Sandlow, R. Sokol, M. Vernon, E. Widra, Diagnostic evaluation of the infertile male: a committee opinion, Fertil. Steril. 103 (2015) e18–e25, http://dx.doi.org/10.1016/j.fertnstert.2014.12.103.
[3] C. Krausz, A.R. Escamilla, C. Chianese, Genetics of male infertility: from research to clinic, Reproduction 150 (2015) R159–R174, http://dx.doi.org/10.1530/REP-15-0261.
[4] A. Amaral, J. Castillo, J. Ramalho-Santos, R. Oliva, The combined human sperm proteome: cellular pathways and implications for basic and clinical science, Hum. Reprod. Update 20 (2014) 40–62, http://dx.doi.org/10.1093/humupd/dmt046.
[5] M. Jodar, E. Sendler, S.I. Moskovtsev, C.L. Librach, R. Goodrich, S. Swanson, R. Hauser, M.P. Diamond, S.a. Krawetz, Absence of sperm RNA elements correlates with idiopathic male infertility, Sci. Transl. Med. 7 (2015), 295re6, http://dx.doi.org/10.1126/scitranslmed.aab1287.
[6] M. Jodar, E. Sendler, S.A. Krawetz, The protein and transcript profiles of human semen, Cell Tissue Res. 363 (2016) 85–96, http://dx.doi.org/10.1007/s00441-015-2237-1.
[7] R.S. Tavares, S. Escada-Rebelo, M. Correia, P.C. Mota, J. Ramalho-Santos, The non-genomic effects of endocrine-disrupting chemicals on mammalian sperm, Reproduction 151 (2016) R1–R13, http://dx.doi.org/10.1530/REP-15-0355.
[8] L. Pantano, M. Jodar, M. Bak, J.L. Ballescà, N. Tommerup, R. Oliva, T. Vavouri, The small RNA content of human sperm reveals pseudogene-derived piRNAs complementary to protein-coding genes, RNA 21 (2015) 1085–1095, http://dx.doi.org/10.1261/rna.046482.114.RNA.
[9] D.T. Carrell, K.I. Aston, R. Oliva, B.R. Emery, C.J. Jonge, The "omics" of human male infertility: integrating big data in a systems biology approach, Cell Tissue Res. 363 (2015) 295–312, http://dx.doi.org/10.1007/s00441-015-2320-7.
[10] R.P. Amann, The cycle of the seminiferous epithelium in humans: a need to revisit? J. Androl. 29 (2008) 469–487, http://dx.doi.org/10.2164/jandrol.107.004655.
[11] F.T.L. Neto, P.V. Bach, B.B. Najari, P.S. Li, M. Goldstein, Spermatogenesis in Humans and Its Affecting Factors, Semin. Cell Dev, Biol, 2016, http://dx.doi.org/10.1016/j.semcdb.2016.04.009.
[12] R. Oliva, G. Dixon, Protamine genes and the histone to protamine replacement reaction, Prog. Nucleic Acid Res. Mol. Biol. 40 (1991) 25–94.
[13] R. Oliva, Protamines and male infertility, Hum. Reprod. Update 12 (2006) 417–435, http://dx.doi.org/10.1093/humupd/dml009.
[14] R. Balhorn, The protamine family of sperm nuclear proteins, Genome Biol. 8 (2007) 227, http://dx.doi.org/10.1186/gb-2007-8-9-227.
[15] M. Jodar, J. Oriola, G. Mestre, J. Castillo, a. Giwercman, J.M. Vidal-Taboada, J.L. Ballescà, R. Oliva, Polymorphisms, haplotypes and mutations in the protamine 1 and 2 genes, Int. J. Androl. 34 (2011) 470–485, http://dx.doi.org/10.1111/j.1365-2605.2010.01115.x.
[16] M. Jodar, R. Oliva, in: E. Baldi, M. Muratori (Eds.), Protamine Alterations in Human Spermatozoa BT - Genetic Damage in Human Spermatozoa, Springer New York, New York, NY 2014, pp. 83–102, http://dx.doi.org/10.1007/978-1-4614-7783-9_6.
[17] T.G. Cooper, Cytoplasmic droplets: the good, the bad or just confusing? Hum. Reprod. 20 (2005) 9–11, http://dx.doi.org/10.1093/humrep/deh555.
[18] R. Sullivan, F. Saez, J. Girouard, G. Frenette, Role of exosomes in sperm maturation during the transit along the male reproductive tract, blood cells, Mol. Dis. 35 (2005) 1–10, http://dx.doi.org/10.1016/j.bcmd.2005.03.005.
[19] I. Björkgren, H. Gylling, H. Turunen, I. Huhtaniemi, L. Strauss, M. Poutanen, P. Sipilä, Imbalanced lipid homeostasis in the conditional Dicer1 knockout mouse epididymis causes instability of the sperm membrane, FASEB J. (2014), http://dx.doi.org/10.1096/fj.14-259382.
[20] U. Sharma, C.C. Conine, J.M. Shea, A. Boskovic, A.G. Derr, X.Y. Bing, C. Belleannee, A. Kucukural, R.W. Serra, F. Sun, L. Song, B.R. Carone, E.P. Ricci, X.Z. Li, L. Fauquier, M.J. Moore, R. Sullivan, C.C. Mello, M. Garber, O.J. Rando, Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals, Science 80− (2015), science.aad6780–, http://dx.doi.org/10.1126/science.aad6780.
[21] T.S. Acott, D.W. Carr, Inhibition of bovine spermatozoa by caudal epididymal fluid: II. Interaction of pH and a quiescence factor, Biol. Reprod. 30 (1984) 926–935, http://dx.doi.org/10.1095/biolreprod30.4.926.
[22] G. Sahlén, O. Nilsson, A. Larsson, L. Carlsson, B.J. Norlén, G. Ronquist, Secretions from seminal vesicles lack characteristic markers for prostasomes, Ups. J. Med. Sci. 115 (2010) 107–112, http://dx.doi.org/10.3109/03009730903366067.
[23] B.F. da Silva, C. Meng, D. Helm, F. Pachl, J. Schiller, E. Ibrahim, C.M. Lynne, N.L. Brackett, R.P. Bertolla, B. Kuster, Towards understanding male infertility after spinal cord injury using quantitative proteomics, Mol. Cell. Proteomics 15 (2016) 1424–1434, http://dx.doi.org/10.1074/mcp.M115.052175.
[24] M. Aalberts, T.A.E. Stout, W. Stoorvogel, Prostasomes: extracellular vesicles from the prostate, Reproduction 147 (2014), http://dx.doi.org/10.1530/REP-13-0358.
[25] A.J. Travis, G.S. Kopf, The role of cholesterol efflux in regulating the fertilization potential of mammalian spermatozoa, J. Clin. Invest. 110 (2002) 731–736, http://dx.doi.org/10.1172/JCI200216392.
[26] L. Vojtech, S. Woo, S. Hughes, C. Levy, L. Ballweber, R.P. Sauteraud, J. Strobl, K. Westerberg, R. Gottardo, M. Tewari, F. Hladik, Exosomes in human semen carry a distinctive repertoire of small non-coding RNAs with potential regulatory functions, Nucleic Acids Res. 42 (2014) 7290–7304, http://dx.doi.org/10.1093/nar/gku347.
[27] J. Castillo, A. Amaral, R. Azpiazu, T. Vavouri, J.M. Estanyol, J.L. Ballescà, R. Oliva, Genomic and proteomic dissection and characterization of the human sperm chromatin, Mol. Hum. Reprod. 20 (2014) 1041–1053, http://dx.doi.org/10.1093/molehr/gau079.
[28] A. Amaral, J. Castillo, J.M. Estanyol, J.L. Ballesca, J. Ramalho-Santos, R. Oliva, Human sperm tail proteome suggests new endogenous metabolic pathways, Mol. Cell. Proteomics 12 (2013) 330–342, http://dx.doi.org/10.1074/mcp.M112.020552.
[29] S. Mateo, J.M. Estanyol, R. Oliva, in: T.D. Carrell, I.K. Aston (Eds.), Spermatogenesis: Methods and Protocols, Humana Press, Totowa, NJ 2013, pp. 411–422, http://dx.doi.org/10.1007/978-1-62703-038-0_35.

[30] A.G. Utleg, E.C. Yi, T. Xie, P. Shannon, J.T. White, D.R. Goodlett, L. Hood, B. Lin, Proteomic analysis of human prostasomes, Prostate 56 (2003) 150–161, http://dx.doi.org/10.1002/pros.10255.
[31] R. Sullivan, F. Saez, Epididymosomes, prostasomes and liposomes; their role in mammalian male reproductive physiology, Reproduction 146 (2013) R21–R35, http://dx.doi.org/10.1530/REP-13-0058.
[32] V. Thimon, G. Frenette, F. Saez, M. Thabet, R. Sullivan, Protein composition of human epididymosomes collected during surgical vasectomy reversal: a proteomic and genomic approach, Hum. Reprod. 23 (2008) 1698–1707, http://dx.doi.org/10.1093/humrep/den181.
[33] R. Oliva, J. Martínez-Heredia, J.M. Estanyol, Proteomics in the study of the sperm cell composition, differentiation and function, Syst. Biol. Reprod. Med. 54 (2008) 23–36, http://dx.doi.org/10.1080/19396360701879595.
[34] R. Oliva, S. de Mateo, J.M. Estanyol, Sperm cell proteomics, Proteomics 9 (2009) 1004–1017, http://dx.doi.org/10.1002/pmic.200800588.
[35] M. Codina, J.M. Estanyol, M.J. Fidalgo, J.L. Ballescà, R. Oliva, Advances in sperm proteomics: best-practise methodology and clinical potential, Expert Rev. Proteomics 12 (2015) 255–277, http://dx.doi.org/10.1586/14789450.2015.1040769.
[36] M. Jodar, S. Selvaraju, E. Sendler, M.P. Diamond, S.a. Krawetz, The presence, role and clinical use of spermatozoal RNAs, Hum. Reprod. Update 19 (2013) 604–624, http://dx.doi.org/10.1093/humupd/dmt031.
[37] M.A. Baker, N. Naumovski, L. Hetherington, A. Weinberg, T. Velkov, R.J. Aitken, Head and flagella subcompartmental proteomic analysis of human spermatozoa, Proteomics 13 (2013) 61–74, http://dx.doi.org/10.1002/pmic.201200350.
[38] S. de Mateo, J. Castillo, J.M. Estanyol, J.L. Ballescà, R. Oliva, Proteomic characterization of the human sperm nucleus, Proteomics 11 (2011) 2714–2726, http://dx.doi.org/10.1002/pmic.201000799.
[39] Y.H. Kim, G. Haidl, M. Schaefer, U. Egner, A. Mandal, J.C. Herr, Compartmentalization of a unique ADP/ATP carrier protein SFEC (sperm flagellar energy carrier, AAC4) with glycolytic enzymes in the fibrous sheath of the human sperm flagellar principal piece, Dev. Biol. 302 (2007) 463–476, http://dx.doi.org/10.1016/j.ydbio.2006.10.004.
[40] D.W. Huang, R.a. Lempicki, B.T. Sherman, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (2009) 44–57, http://dx.doi.org/10.1038/nprot.2008.211.
[41] D.W. Huang, B.T. Sherman, R.a. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, Nucleic Acids Res. 37 (2009) 1–13, http://dx.doi.org/10.1093/nar/gkn923.
[42] J. Castillo, J. Estanyol, J. Ballescà, R. Oliva, Human sperm chromatin epigenetic potential: genomics, proteomics, and male infertility, Asian J. Androl. 17 (2015) 601–609, http://dx.doi.org/10.4103/1008-682X.153302.
[43] C. González-marín, J. Gosálvez, R. Roy, Types, causes, detection and repair of DNA fragmentation in animal and human sperm cells, Int. J. Mol. Sci. 13 (2012) 14026–14052, http://dx.doi.org/10.3390/ijms131114026.
[44] R. Bose, G. Manku, M. Culty, S.S. Wing, Ubiquitin-proteasome system in spermatogenesis, Adv. Exp. Med. Biol. 759 (2014) 181–213, http://dx.doi.org/10.1007/978-1-4939-0817-2-9.
[45] P. Sutovsky, Sperm proteasome and fertilization, Reproduction 142 (2011) 1–14, http://dx.doi.org/10.1530/REP-11-0041.
[46] C. Hajjar, K.M. Sampuda, L. Boyd, Dual roles for ubiquitination in the processing of sperm organelles after fertilization, BMC Dev. Biol. 14 (2014) 1–10, http://dx.doi.org/10.1186/1471-213X-14-6.
[47] P. Morales, E. Pizarro, M. Kong, M. Jara, Extracellular localization of proteasomes in human sperm, Mol. Reprod. Dev. 68 (2004) 115–124, http://dx.doi.org/10.1002/mrd.20052.
[48] K. Mochida, L.L. Tres, A.L. Kierszenbaum, Structural features of the 26S proteasome complex isolated from rat testis and sperm tail, Mol. Reprod. Dev. 57 (2000) 176–184, http://dx.doi.org/10.1002/1098-2795(200010)57:2<176::AID-MRD9>3.0.CO;2-O.
[49] S.S. du Plessis, A. Agarwal, G. Mohanty, M. van der Linde, Oxidative phosphorylation versus glycolysis: what fuel do spermatozoa use? Asian J. Androl. 17 (2015) 230–235, http://dx.doi.org/10.4103/1008-682X.135123.
[50] K. Gilany, A. Minai-Tehrani, E. Savadi-Shiraz, H. Rezadoost, N. Lakpour, Exploring the human seminal plasma proteome: an unexplored gold mine of biomarker for male infertility and male reproduction disorder, J. Reprod. Infertil. 16 (2015) 61–71.
[51] A.P. Drabovich, P. Saraon, K. Jarvi, E.P. Diamandis, Seminal plasma as a diagnostic fluid for male reproductive system disorders, Nat. Rev. Urol. 11 (2014) 278–288, http://dx.doi.org/10.1038/nrurol.2014.74.
[52] A.D. Rolland, R. Lavigne, C. Dauly, P. Calvel, C. Kervarrec, T. Freour, B. Evrard, N. Rioux-Leclercq, J. Auger, C. Pineau, Identification of genital tract markers in the human seminal plasma using an integrative genomics approach, Hum. Reprod. 28 (2013) 199–209, http://dx.doi.org/10.1093/humrep/des360.
[53] M.A. Lizarbe, J.I. Barrasa, N. Olmo, F. Gavilanes, J. Turnay, Annexin-phospholipid interactions. Functional implications, Int. J. Mol. Sci. 14 (2013) 2652–2683, http://dx.doi.org/10.3390/ijms14022652.
[54] K.G. Ronquist, B. Ek, J. Morrell, A. Stavreus-Evers, B. Ström Holst, P. Humblot, G. Ronquist, A. Larsson, Prostasomes from four different species are able to produce extracellular adenosine triphosphate (ATP), Biochim. Biophys. Acta Gen. Subj. 1830 (2013) 4604–4610, http://dx.doi.org/10.1016/j.bbagen.2013.05.019.
[55] B. Stegmayr, G. Ronquist, Promotive effect on human sperm progressive motility by prostasomes, Urol. Res. 10 (1982) 253–257, http://dx.doi.org/10.1007/BF00255932.
[56] M. Battistone, N. Brukman, G. Carvajal, D. Cohen, P. Cuasnicú, L. Curci, V. Da Ros, M. Gómez Elías, M. Muñoz, From the epididymis to the egg: participation of CRISP proteins in mammalian fertilization, Asian J. Androl. 0 (0) (2015), http://dx.doi.org/10.4103/1008-682X.155769.
[57] G.D. Johnson, P. Mackie, M. Jodar, S. Moskovtsev, S.A. Krawetz, Chromatin and extracellular vesicle associated sperm RNAs, Nucleic Acids Res. 43 (2015) 6847–6859, http://dx.doi.org/10.1093/nar/gkv591.
[58] M. Liu, Z. Hu, L. Qi, J. Wang, T. Zhou, Y. Guo, Y. Zeng, B. Zheng, Y. Wu, P. Zhang, X. Chen, W. Tu, T. Zhang, Q. Zhou, M. Jiang, X. Guo, Z. Zhou, J. Sha, Scanning of novel cancer/testis proteins by human testis proteomic analysis, Proteomics 13 (2013) 1200–1210, http://dx.doi.org/10.1002/pmic.201200489.
[59] G. Manandhar, H. Schatten, P. Sutovsky, Centrosome reduction during gametogenesis and its significance, Biol. Reprod. 72 (2005) 2–13, http://dx.doi.org/10.1095/biolreprod.104.031245.
[60] W. Alasmari, C.L.R. Barratt, S.J. Publicover, K.M. Whalley, E. Foster, V. Kay, S. Martins Da Silva, S.K. Oxenham, The clinical significance of calcium-signalling pathways mediating human sperm hyperactivation, Hum. Reprod. 28 (2013) 866–876, http://dx.doi.org/10.1093/humrep/des467.
[61] F. Saez, A. Ouvrier, J.R. Drevet, Epididymis cholesterol homeostasis and sperm fertilizing ability, Asian J. Androl. 13 (2011) 11–17, http://dx.doi.org/10.1038/aja.2010.64.
[62] B.A. Laflamme, M.F. Wolfner, Identification and function of proteolysis regulators in seminal fluid, Mol. Reprod. Dev. 80 (2013) 80–101, http://dx.doi.org/10.1002/mrd.22130.
[63] E.Y. Kim, E.H. Noh, E.J. Noh, M.J. Park, H.Y. Park, D.S. Lee, K.Z. Riu, S.P. Park, Effect of glycosaminoglycans on in vitro fertilizing ability and in vitro developmental potential of bovine embryos, Asian-Australasian J. Anim. Sci. 26 (2013) 178–188, http://dx.doi.org/10.5713/ajas.2012.12406.
[64] G.R. Dohle, G.M. Colpi, T.B. Hargreave, G.K. Papp, A. Jungwirth, W. Weidner, EAU guidelines on male infertility, Eur. Urol. 48 (2005) 703–711, http://dx.doi.org/10.1016/j.eururo.2005.06.002.
[65] F. Edition, Examination and processing of human semen, World Health. Edition F (2010) 286. http://whqlibdoc.who.int/publications/2010/9789241547789_eng.pdf.
[66] C. Zhao, R. Huo, F.-Q. Wang, M. Lin, Z.-M. Zhou, J.-H. Sha, Identification of several proteins involved in regulation of sperm motility by proteomic analysis, Fertil. Steril. 87 (2007) 436–438, http://dx.doi.org/10.1016/j.fertnstert.2006.06.057.
[67] J. Martínez-Heredia, S. de Mateo, J.M. Vidal-Taboada, J.L. Ballescà, R. Oliva, Identification of proteomic differences in asthenozoospermic sperm samples, Hum. Reprod. 23 (2008) 783–791, http://dx.doi.org/10.1093/humrep/den024.
[68] C.C. Chan, H.A. Shui, C.H. Wu, C.Y. Wang, G.H. Sun, H.M. Chen, G.J. Wu, Motility and protein phosphorylation in healthy and asthenozoospermic sperm, J. Proteome Res. 8 (2009) 5382–5386, http://dx.doi.org/10.1021/pr9003932.
[69] A.B. Siva, D.B. Kameshwari, V. Singh, K. Pavani, C.S. Sundaram, N. Rangaraj, M. Deenadayal, S. Shivaji, Proteomics-based study on asthenozoospermia: differential expression of proteasome alpha complex, Mol. Hum. Reprod. 16 (2010) 452–462, http://dx.doi.org/10.1093/molehr/gaq009.
[70] T. Botta, S. Blescia, J. Martínez-Heredia, R. Lafuente, M. Brassesco, J. Luis Ballescà, R. Oliva, Identificación de diferencias proteómicas en muestras oligozoospérmicas, Rev. Int. Androl. 7 (2009) 14–19, http://dx.doi.org/10.1016/S1698-031X(09)70257-2.
[71] H.C.A. Chao, C.L. Chung, H.A. Pan, P.C. Liao, P.L. Kuo, C.C. Hsu, Protein tyrosine phosphatase non-receptor type 14 is a novel sperm-motility biomarker, J. Assist. Reprod. Genet. 1–11 (2011), http://dx.doi.org/10.1007/s10815-011-9602-0.
[72] S. Shen, J. Wang, J. Liang, D. He, Comparative proteomic study between human normal motility sperm and idiopathic asthenozoospermia, World J. Urol. 31 (2013) 1395–1401, http://dx.doi.org/10.1007/s00345-013-1023-5.
[73] J. Wang, J. Wang, H.-R. Zhang, H.-J. Shi, D. Ma, H.-X. Zhao, B. Lin, R.-S. Li, Proteomic analysis of seminal plasma from asthenozoospermia patients reveals proteins that affect oxidative stress responses and semen quality, Asian J. Androl. 11 (2009) 484–491, http://dx.doi.org/10.1038/aja.2009.26.
[74] E. Giacomini, B. Ura, E. Giolo, S. Luppi, M. Martinelli, R.C. Garcia, G. Ricci, Comparative analysis of the seminal plasma proteomes of oligoasthenozoospermic and normozoospermic men, Reprod. BioMed. Online 30 (2015) 522–531, http://dx.doi.org/10.1016/j.rbmo.2015.01.010.
[75] R. Sharma, a. Agarwal, a.J. Hamada, R. Jesudasan, S. Yadav, E. Sabanegh, Proteomic analysis of seminal plasma proteins in men with various semen parameters, Fertil. Steril. 98 (2012) S148, http://dx.doi.org/10.1016/j.fertnstert.2012.07.547.
[76] T.-T. Liao, Z. Xiang, W.-B. Zhu, L.-Q. Fan, Proteome analysis of round-headed and normal spermatozoa by 2-D fluorescence difference gel electrophoresis and mass spectrometry, Asian J. Androl. 11 (2009) 683–693, http://dx.doi.org/10.1038/aja.2009.59.
[77] S. Thacker, S.P. Yadav, R.K. Sharma, A. Kashou, B. Willard, D. Zhang, A. Agarwal, Evaluation of sperm proteins in infertile men: a proteomic approach, Fertil. Steril. 95 (2011) 2745–2748, http://dx.doi.org/10.1016/j.fertnstert.2011.03.112.
[78] P.P. Parte, P. Rao, S. Redij, V. Lobo, S.J. D'Souza, R. Gajbhiye, V. Kulkarni, Sperm phosphoproteome profiling by ultra performance liquid chromatography followed by data independent analysis (LC-MSE) reveals altered proteomic signatures in asthenozoospermia, J. Proteome 75 (2012) 5861–5871, http://dx.doi.org/10.1016/j.jprot.2012.07.003.
[79] A. Amaral, C. Paiva, C. Attardo Parrinello, J.M. Estanyol, J.L. Ballescà, J.J. Ramalho-Santos, R. Oliva, Identification of proteins involved in human sperm motility using high-throughput differential proteomics, J. Proteome Res. 13 (2014) 5670–5684, http://dx.doi.org/10.1021/pr500652y.
[80] R. Bhilawadikar, K. Zaveri, L. Mukadam, S. Naik, K. Kamble, D. Modi, I. Hinduja, Levels of Tektin 2 and CatSper 2 in normozoospermic and oligoasthenozoospermic men and its association with motility, fertilization rate, embryo quality and

pregnancy rate, J. Assist. Reprod. Genet. 1–11 (2013), http://dx.doi.org/10.1007/s10815-013-9972-6.
[81] E. Salvolini, E. Buldreghini, G. Lucarini, A. Vignini, A. Giulietti, A. Lenzi, L. Mazzanti, R. Di Primio, G. Balercia, Interleukin-1β, cyclooxygenase-2, and hypoxia-inducible factor-1α in asthenozoospermia, Histochem. Cell Biol. 142 (2014) 569–575, http://dx.doi.org/10.1007/s00418-014-1232-z.
[82] H. Li, N. Yu, X. Zhang, W. Jin, H. Li, Spermatozoal protein profiles in male infertility with asthenozoospermia, Chin. Med. J. (Engl). (2010) 2879–2882.
[83] C.N. An, H. Jiang, Q. Wang, R.P. Yuan, J.M. Liu, W.L. Shi, Z.Y. Zhang, X.P. Pu, Down-regulation of DJ-1 protein in the ejaculated spermatozoa from Chinese asthenozoospermia patients, Fertil. Steril. 96 (2011) 19–23, e2, http://dx.doi.org/10.1016/j.fertnstert.2011.04.048.
[84] W. Sun, Q. Guan, J. Wen, Q. Zhang, W. Yang, B. Zhang, W. Cui, Z. Zou, Y. Yu, Calcium- and integrin-binding protein-1 is down-regulated in the sperm of patients with oligoasthenozoospermia: CIB1 expression in patients with oligoasthenozoospermia, J. Assist. Reprod. Genet. 31 (2014) 541–547, http://dx.doi.org/10.1007/s10815-014-0177-4.
[85] M. Motiei, M. Tavalaee, F. Rabiei, R. Hajihosseini, M.H. Nasr-Esfahani, Evaluation of HSPA2 in fertile and infertile individuals, Andrologia 45 (2013) 66–72, http://dx.doi.org/10.1111/j.1439-0272.2012.01315.x.
[86] Y.C. Tasi, H.C.A. Chao, C.L. Chung, X.Y. Liu, Y.M. Lin, P.C. Liao, H.A. Pan, H.S. Chiang, P.L. Kuo, Y.H. Lin, Characterization of 3-hydroxyisobutyrate dehydrogenase, HIBADH, as a sperm-motility marker, J. Assist. Reprod. Genet. 30 (2013) 1–8, http://dx.doi.org/10.1007/s10815-013-9954-8.
[87] Y. Liu, G. Chen, L. Lu, H. Sun, Q. Guo, K. Xue, Y. Fan, Z. Ding, RNASET2 in human spermatozoa and seminal plasma: a novel relevant indicator for asthenozoospermia, Andrology 1 (2013) 75–84, http://dx.doi.org/10.1111/j.2047-2927.2012.00022.x.
[88] R. Tomar, A.K. Mishra, N.K. Mohanty, A.K. Jain, Altered expression of succinic dehydrogenase in asthenozoospermia infertile male, Am. J. Reprod. Immunol. 68 (2012) 486–490, http://dx.doi.org/10.1111/aji.12023.
[89] X. Zhang, R. Diao, X. Zhu, Z. Li, Z. Cai, Metabolic characterization of asthenozoospermia using nontargeted seminal plasma metabolomics, Clin. Chim. Acta 450 (2015) 254–261, http://dx.doi.org/10.1016/j.cca.2015.09.001.
[90] K.K. Shukla, S. Agnihotri, A. Gupta, A.A. Mahdi, E.A. Mohamed, S.N. Sankhwar, P. Sharma, Significant association of TNFα and IL-6 gene with male infertility-an explorative study in Indian populations of Uttar Pradesh, Immunol. Lett. 156 (2013) 30–37, http://dx.doi.org/10.1016/j.imlet.2013.08.011.
[91] H. Margaryan, A. Dorosh, J. Capkova, P. Manaskova-Postlerova, A. Philimonenko, P. Hozak, J. Peknicova, Characterization and possible function of glyceraldehyde-3-phosphate dehydrogenase-spermatogenic protein GAPDHS in mammalian sperm, Reprod. Biol. Endocrinol. 13 (2015) 15, http://dx.doi.org/10.1186/s12958-015-0008-1.
[92] M. Luconi, G. Cantini, E. Baldi, G. Forti, Role of a-kinase anchoring proteins (AKAPs) in reproduction, Front. Biosci. 16 (2011) 1315–1330, http://dx.doi.org/10.2741/3791.
[93] H. Tarnasky, M. Cheng, Y. Ou, J.C. Thundathil, R. Oko, F.a. van der Hoorn, Gene trap mutation of murine outer dense fiber protein-2 gene can result in sperm tail abnormalities in mice with high percentage chimaerism, BMC Dev. Biol. 10 (2010) 67, http://dx.doi.org/10.1186/1471-213X-10-67.
[94] S. Hansen, E.E. Eichler, S.M. Fullerton, D. Carrell, SPANX gene variation in fertile and infertile males, Syst. Biol. Reprod. Med. 55 (2010) 18–26, http://dx.doi.org/10.3109/19396360903312015.
[95] R. Azpiazu, A. Amaral, J. Castillo, J.M. Estanyol, M. Guimer, J.L. Ballescà, J. Balasch, R. Oliva, High-throughput sperm differential proteomics suggests that epigenetic alterations contribute to failed assisted reproduction, Hum. Reprod. 29 (2014) 1225–1237, http://dx.doi.org/10.1093/humrep/deu073.
[96] W. Xu, H. Hu, Z. Wang, X. Chen, F. Yang, Z. Zhu, P. Fang, J. Dai, L. Wang, H. Shi, Z. Li, Z. Qiao, Proteomic characteristics of spermatozoa in normozoospermic patients with infertility, J. Proteome 75 (2012) 5426–5436, http://dx.doi.org/10.1016/j.jprot.2012.06.021.
[97] S. McReynolds, M. Dzieciatkowska, J. Stevens, K.C. Hansen, W.B. Schoolcraft, M.G. Katz-Jaffe, Toward the identification of a subset of unexplained infertility: a sperm proteomic approach, Fertil. Steril. 102 (2014) 692–699, http://dx.doi.org/10.1016/j.fertnstert.2014.05.021.
[98] Y. Zhu, Y. Wu, K. Jin, H. Lu, F. Liu, Y. Guo, F. Yan, W. Shi, Y. Liu, X. Cao, H. Hu, H. Zhu, X. Guo, J. Sha, Z. Li, Z. Zhou, Differential proteomic profiling in human spermatozoa that did or did not result in pregnancy via IVF and AID, Proteomics Clin. Appl. 7 (2013) 850–858, http://dx.doi.org/10.1002/prca.201200078.
[99] A.P. Drabovich, A. Dimitromanolakis, P. Saraon, A. Soosaipillai, I. Batruch, B. Mullen, K. Jarvi, E.P. Diamandis, Differential diagnosis of azoospermia with proteomic biomarkers ECM1 and TEX101 quantified in seminal plasma, Sci. Transl. Med. 5 (2013), 212ra160, http://dx.doi.org/10.1126/scitranslmed.3006260.
[100] A. Agarwal, R. Sharma, D. Durairajanayagam, Z. Cui, A. Ayaz, S. Gupta, B. Willard, B. Gopalan, E. Sabanegh, Differential proteomic profiling of spermatozoal proteins of infertile men with unilateral or bilateral varicocele, Urology 85 (2015) 580–588, http://dx.doi.org/10.1016/j.urology.2014.11.030.
[101] A. Agarwal, R. Sharma, L. Samanta, D. Durairajanayagam, E. Sabanegh, Proteomic signatures of infertile men with clinical varicocele and their validation studies reveal mitochondrial dysfunction leading to infertility, Asian J. Androl. 18 (2016) 282–291, http://dx.doi.org/10.4103/1008-682X.170445.
[102] P.T. Del Giudice, L.B. Belardin, M. Camargo, D.S. Zylbersztejn, V.M. Carvalho, K.H.M. Cardozo, R.P. Bertolla, A.P. Cedenho, Determination of testicular function in adolescents with varicocoele - a proteomics approach, Andrology 4 (2016) 447–455, http://dx.doi.org/10.1111/andr.12174.

7.3. Annex 3

Funció mitocondrial alterada en espermatozoides de pacients amb fallida de fertilització repetitiva després de ICSI revelada per proteòmica.

En preparació

Altered mitochondrial function in sperm from patients with repetitive fertilization failure after ICSI revealed by proteomics.
Torra-Massana M, Jodar M, Barragán M, Soler-Ventura A, Delgado-Dueñas D, Rodríguez A, Oliva R*, Vassena R*

Article in preparation

Altered mitochondrial function in sperm from patients with repetitive fertilization failure after ICSI revealed by proteomics

Marc Torra-Massana[1,2], Meritxell Jodar[2,3], Montserrat Barragán[1], Ada Soler-Ventura[2,3], David Delgado-Dueñas[2,3], Amelia Rodríguez[1], Rafael Oliva[2,3*], Rita Vassena[1*]

[1]EUGIN, Balmes 236, 08006, Barcelona, Spain. [2]Molecular Biology of Reproduction and Development Group, Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS), Fundació Clínic per a la Recerca Biomèdica, Hospital Clinic, Faculty of Medicine, University of Barcelona, Casanova 143, 08036 Barcelona, Spain. [3]EUGIN-UB Research Excellence Program
*Correspondence: rvassena@eugin.es; roliva@ub.edu

ABSTRACT:

Unexplained fertilization failure (FF), occurring in 1-3% of intracytoplasmic sperm injection (ICSI) cycles, results in both psychological and financial burden for the patients. However, the molecular causes behind FF remain largely unknown. Mass spectrometry is a powerful technique to identify and quantify proteins across samples; however, no study so far has used it to dissect the proteomic signature of sperm with FF after ICSI. The aim of this study was to investigate whether sperm samples from patients suffering repetitive FF after ICSI display alterations in their protein content. To that end, 17 infertile men were included: 5 patients presented FF in ≥3 consecutive ICSI cycles, while 12 patients had a fertilization rate >75% (controls). Individual sperm samples were subjected to 2D-LC-MS/MS. Both conventional and novel statistical approaches were used to identify differentially abundant proteins. Additionally, analysis of mitochondrial and - proteasomal abundance and activity were performed, using western blot, FACS analysis of JC-1 staining and AMC-peptide fluorometric assay. Four proteins presented lower abundance (FMR1NB, FAM209B, RAB2B, PSMA1) in the FF group compared to controls, while five mitochondrial proteins presented higher abundance in FF (DLAT, ATP5H, SLC25A3, SLC25A6, FH) ($p<0.05$). The altered abundance of mitochondrial DLAT and proteasomal PSMA1 was corroborated by Western-Blot. Of relevance, novel stable-protein pair analysis identified 73 correlations comprising 28 proteins within controls, while different mitochondrial proteins (i.e. PDHA2, PHB2, ATP5F1D) lost >50% of these correlations in specific FF samples pointing out

specific mitochondrial deregulations. This is the first proteomic analysis of sperm from patients who resulted in fertilization failure after ICSI. The altered proteins, most of them related to mitochondrial function, could help to identify diagnostic/prognostic markers of fertilization failure and could further dissect the molecular paternal contribution to reach successful fertilization. In conclusion, sperm samples from patients with FF after ICSI present altered abundance of different proteins, including mainly mitochondrial proteins.

INTRODUCTION

Intracytoplasmic sperm injection (ICSI) has become the most commonly used assisted reproduction technique in fertility clinics from Europe and USA, including patients with non-severe male factor infertility (Kupka et al. 2014, Boulet et al. 2015). ICSI is very efficient, with a mean fertilization rate around 70-80%; however, total fertilization failure (TFF) still occurs in 1-3% of cycles, even when inseminating an appropriate number of oocytes (Flaherty et al. 1998, Yanagida 2004). Total or low fertilization failure (FF) cannot be easily diagnosed or predicted, resulting in difficult clinical management and counseling of these patients, along with high economic and emotional burden. Since ICSI bypasses most of the biological barriers involved in fertilization, the failure to induce oocyte activation has been proposed as the most probable cause of FF after ICSI (Flaherty et al. 1998). For instance, it is well established that calcium oscillations are essential at the onset of oocyte activation (Yeste et al. 2016). However, little is known about other molecular mechanisms involved in the fertilization process beyond sperm-oocyte fusion such as sperm decondensation, polyspermy block, resumption of meiosis, formation of the pronuclei and generation of the first mitotic spindle (Tosti and Menezo 2016).

Sperm defects such as poor nuclear chromatin condensation, centrosomal dysfunction, or deficient sperm head–tail attachment may lead to diminished sperm fertilizing ability (Nasr-Esfahani et al. 2007, Rawe et al. 2008, Terada et al. 2010). Indeed, it has been demonstrated that oocyte activation is triggered by sperm-borne factors, and therefore the male factor should explain some of the FF after ICSI (Yeste et al. 2016). For instance, mutations in sperm phospholipase C zeta gene (*PLCZ1*) have been detected in patients experiencing oocyte activation failure (Kashir et al. 2018, Torra-Massana et al. 2019). However, knock-out mice for *PLCZ1* are still able to generate offspring, suggesting that additional sperm factors are required to achieve a successful fertilization (Nozawa et al. 2018). This hypothesis is also supported by recent findings of our group showing that around half of the patients presenting FF after ICSI do not carry missense variants on *PLCZ1*, nor show alterations in their protein content or localization (Ferrer-Vaquer et al. 2016, Torra-Massana et al. 2019). Therefore, other currently unknown sperm factors could play important roles during the early fertilization events after ICSI.

Up to date, around 7,000 proteins have been identified using high-throughput techniques in sperm (compiled in (Castillo et al. 2018)). In addition, comparative proteomics analyses have emerged as a useful approach to understand and characterize the mechanisms underlying male infertility (reviewed in (Jodar 2017)). Alterations in the sperm proteome have been analyzed in different infertile phenotypes, as for example, asthenozoospermia (Amaral et al. 2014), globozoospermia (Liao et al. 2009), high oxidative stress (Sharma et al. 2013), high DNA fragmentation (Intasqui et al. 2013) or varicocele (Fariello et al. 2012). So far, four studies have addressed proteomic alterations in sperm samples from patients with FF after conventional IVF, identifying proteins with differential abundance mainly related to sperm motility and gamete interaction (Pixton et al. 2004, Frapsauce et al. 2014, Légaré 2014, Liu et al. 2018). However, to the best of our knowledge, sperm factors associated with fertilization failure after ICSI have not been previously assessed by proteomics-based approaches.

In this study we aimed to identify novel male factors associated with defects in oocyte activation and early fertilization events by comparing individual sperm samples from patients with good fertilization rate (FR >75%) and repetitive FF (FR < 20 %) using high-throughput proteomics approaches (LC-MS/MS). Some of the proteins and mechanisms identified as altered in this study could be useful as clinical diagnostic and prognostic markers for FF after ICSI.

MATERIALS AND METHODS

Ethical approval

This study has been done in accordance with the ethical principles stated in the Declaration of Helsinki (1964), as revised in 2013. Approval to conduct this research was obtained from the local Ethical Committee for Clinical Research (CEIm) of Clinica Eugin before the beginning; written consent was obtained from all patients enrolled in the study.

Study population

Patients were recruited from January 2016 to January 2018. Men above 45 years old, with <15 % sperm progressive motility or <10 million/ml sperm count, and couples with ICSI cycles using oocytes from women older than 42 years old were excluded. A total of 17 couples was included in this study, all with similar clinical characteristics (Table 1 and Table S1). Patients with an overall fertilization rate above 75 % after ICSI were included in the control group (n = 12, C1-C12). Patients who experienced repetitive FF or poor FR in at least 3 consecutive ICSI cycles (overall FR < 20 %), and without any previous IVF

cycle with good fertilization rate, were included in the FF group (n= 5, F1-F5). A total of at least 6 oocytes were inseminated per patient.

For proteomic approaches and western blot validation, 8 controls and 4 FF samples were used. After proteomic analysis was performed, additional samples (one for FF group and 4 for control group) were included (F5, C9-C12). For functional characterization of mitochondrial and proteasomal function, 5 controls and 4 FF samples were included (as shown in Table S2), while sample F2 could not be used to this aim due to insufficient amount of protein extract.

Sperm sample collection

Semen samples were obtained in a sterile container by masturbation after four days of abstinence, at mean (3 to 10 days; mean ± SD; 4.17 ± 1.74). After 30 minutes of liquefaction at room temperature, sperm samples were analyzed with an integrated semen analysis system (ISAS, Proiser, Paterna, Spain), and classified following standard criteria (WHO. 2010). The ejaculate volume left after insemination by ICSI was cryopreserved the same day in either straws or cryotubes using Sperm Cryoprotect II (Nidacon International AB, Gothenburg, Sweden) according to manufacturer's instructions, and stored at -80 ºC until further processing.

Ovarian stimulation and ICSI procedure

Ovarian stimulation was based on either long agonist or antagonist protocol depending on medical history, clinical characteristics, and ovarian reserve of the patient, while short antagonist and GnRH agonist triggers were used for oocyte donors. ICSI was performed following standard procedures on metaphase-II oocytes (Palermo et al. 1992). Fertilization was assessed between 16 to 19 hours post-ICSI by the presence or absence of 2 pronuclei (PN) in the zygote.

Sperm processing and protein solubilization

The general strategy used in this study is depicted in Fig. 1. Sperm samples frozen on straws and cryotubes were thawed for 5 minutes at 37 ºC or for 3 minutes in a water bath at 40 ºC, respectively. Once thawed, sperm samples were centrifuged at room temperature for 10 minutes at 500 g. Each sperm pellet was suspended in 1 ml PureSperm® Buffer (Nidacon) and assessed for both sperm concentration and round cell contamination. Sperm samples were then purified by density gradient centrifugation (DGC) in 50 % PureSperm® (Nidacon) following manufacturer's instructions. After DGC, the recovered fraction (pellet) contained less than 1 % of contaminating round cells. Samples were pelleted at 2,500 g for 5 minutes, and 25 µl of lysis buffer (2% SDS, 1 mM

phenylmethylsulphonyl fluoride (PMSF), in PBS) were added to resuspend and lyse 10 million sperm. Protein extracts were stored at -80 ºC until processing for proteomics and Western Blot.

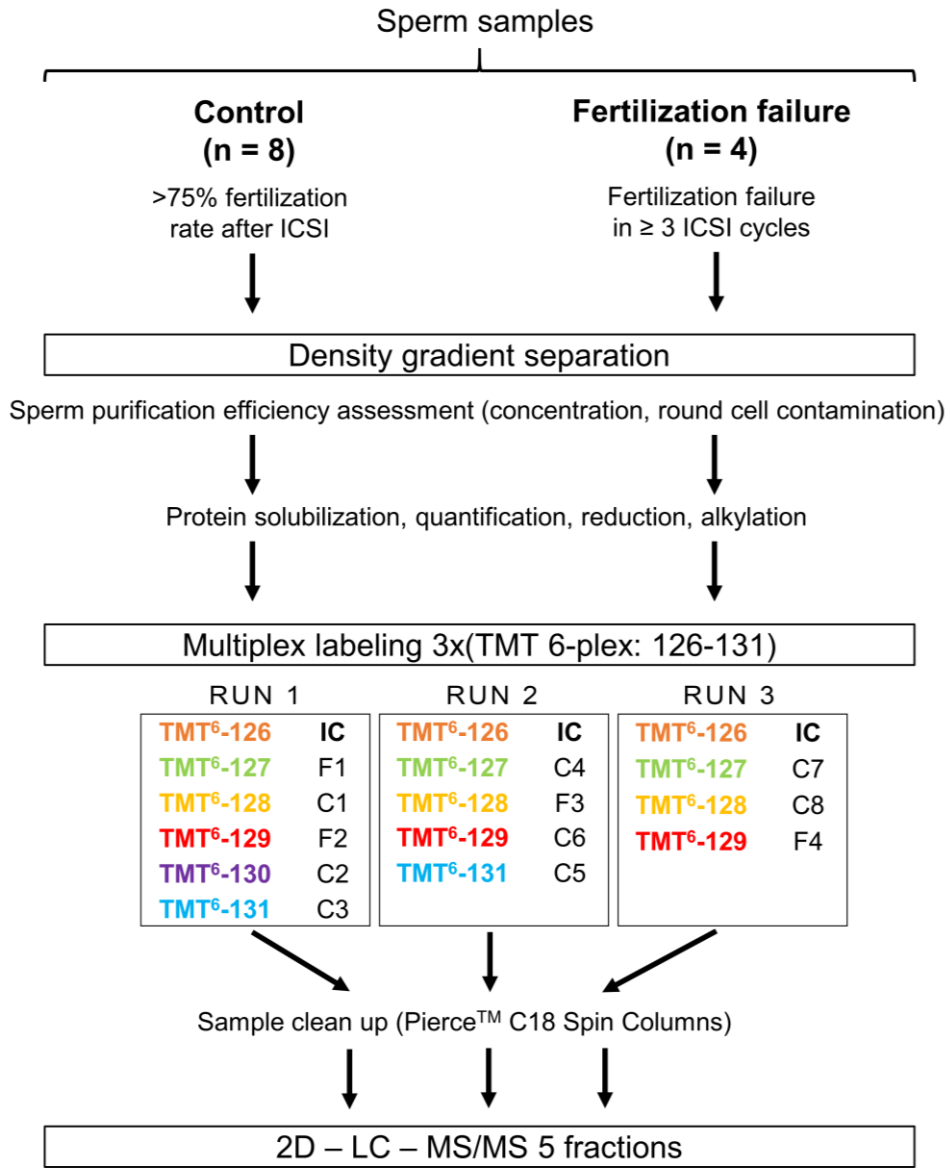


Figure 1. Schematic representation of the strategy used for the identification of proteomic changes between study groups (controls and FF). Samples were thawed and processed by density gradient in 50 % PureSperm®. After quality assessment (concentration and round cell contamination), cells were lysed and soluble proteins were quantified, digested, reduced, alkylated, and labeled with TMT 6-plex isobaric tags as shown in Material and methods section. Prior to labelling, equal amounts of peptides from each sample (including FF and controls) were pooled as internal control sample (IC) and loaded in all three independent multiplex runs (RUN 1, 2, and 3). Labelled peptides were desalted and cleaned up before being separated by liquid chromatography, 5 fractions were collected, and proteins were identified and quantified through MS/MS.

Tandem Mass Tag (TMT) labelling

Protein extracts were thawed at room temperature for 10 minutes followed by 30 minutes at 4 ºC with continuous shaking, and total protein abundance was quantified using microBCA™ Protein Assay Kit (Thermo Fisher Scientific, Rockford, IL, USA) following the manufacturer's instructions. Subsequently, TMT-6-plex (Thermo Fisher Scientific) labeling was applied, following the manufacturer's instructions with minor modifications. Briefly, 25 µg of protein extract from each sample were diluted in 100 mM triethyl ammonium bicarbonate (TEAB; pH 8.5, Thermo Fisher Scientific) to reach a final protein concentration of 0.6 µg/µl. Afterwards, proteins were reduced in 9.5 mM tris(2-carboxyethyl) phosphine (TCEP) for one hour at 55 °C, and alkylated in a solution containing a final concentration of 17 mM iodoacetamide for 30 minutes in the dark. Proteins were precipitated overnight at -20 ºC by adding cold 100 % acetone to a final percentage of 87.5 %. Samples were centrifuged at 4ºC for 10 minutes at 17,500 g and precipitated proteins were dissolved in 50 mM TEAB to obtain a protein concentration of 0.6 µg/µl. Protein trypsinization was performed overnight at 37ºC with continuous shaking, in 1:20 protease-to-protein ratio. Prior to labeling, one fifth of the processed sample (equivalent to 4.8 µg) was removed from each sample and all combined, to establish the internal control sample (IC) used to compare the results of the three different multiplexes (Fig. 1). Samples and internal control were incubated for 1 hour with TMT-6-plex reagents previously dissolved in anhydrous acetonitrile and reactions were quenched by incubating for 15 minutes in a final concentration of 0.29 % hydroxylamine solution. TMT-labeled samples were combined at equivalent amounts (10 µg per sample) within each multiplex, each one containing one IC, and at least one FF sample and 2 controls (Fig. 1). Combined and labeled peptides were dried in a speed-vacuum centrifuge and dissolved in a solution containing 0.5 % trifluoroacetic acid (TFA) and 5 % acetonitrile. Afterwards, TMT-labeled mixtures were desalted by using reversed-phase C18 Spin Columns (Pierce C18 Spin Columns, Thermo Fisher Scientific) following manufacturer's instructions, and eluted in 40 µl of 70 % acetonitrile.

Strong-cation exchange (SCX) chromatography fractionation

TMT-labeled peptides from each multiplex were fractionated using an Oasis MCX 1 cc Vac Cartridge (10 mg Sorbent per Cartridge, 30 µm Particle Size, Waters, Milford, MA, USA). Initially, the sorbent was activated by flowing through 1 ml of 100 % acetonitrile and equilibrated with a solution (EqS) containing 5 mM ammonium formate (AF) and 25 % acetonitrile (ACN) (pH = 3). 5 µl of each multiplex was loaded to the column and, after washing twice with EqS, sequential elution steps were performed by gravity using a gradient of ammonium formate solutions (200 mM AF / 25 % CAN; 350 mM AF / 25 %

CAN; 1000 mM AF / 25 % CAN; 500 mM AF / 25 % ACN / 1500 mM KCl; and 1000 mM AF / 50 % ACN). The collected fractions (5 fractions per multiplex, 15 fractions in total) were desalted using a reversed-phase C18 Spin Columns (Pierce C18 Spin Columns, Thermo Fisher Scientific) following manufacturer's instructions.

Protein identification and quantification by liquid chromatography - tandem mass spectrometry (LC-MS/MS)

Desalted fractions were brought to dryness under a speed-vacuum and reconstituted with 5 µl of 0.1 % formic acid and 4.5 µl from each fraction was loaded onto a trap column C18 (2 cm length (L), 200 µm inner diameter (ID), 5 µm particle size, 120 Å pore size; Nanoseparations, Nieuwkoop, Netherlands). For separation, a linear gradient was applied together with the analytical column (15 cm L, 75 µm ID, 3 µm particle size, 100 Å pore size, Thermo Fisher Scientific). Buffer A (97 % H_2O, 3 % acetonitrile, 0.1 % formic acid) and buffer B (3 % H_2O, 97 % acetonitrile, 0.1 % formic acid) were used for a 80 minutes linear gradient as follows (0-5 minutes 0 % to 0 % B, 5-60 minutes 0 % to 37 % B, 60-65 minutes 37 % to 100 % B, 65-80 minutes 100 % to 100 % B) for elution at a rate of 400 nl / minute to separate the peptides. The MS/MS analysis was performed by a nano-LC ultra 2D eksigent (AB Sciex,Baden, Switzerland) attached to a LTQ-OrbitrapVelos coupled to a nanospray ion source (Thermo Fisher Scientific). The LTQ OrbitrapVelos (Thermo Fisher Scientific) settings included 30,000 of resolution for the MS1 scans at 400 m/z for precursor ions followed by 7,500 of resolution for the MS2 scans of the 15 most intense precursor ions at 400 m/z, in positive ion mode. The mass range was set at 380-1,500 m/z. The acquisition of MS/MS data was performed using Xcalibur 2.2 (Thermo Fisher Scientific). Normalized collision energy for higher-energy collisional dissociation (HCD)-MS2 was set to 42 %.

LC-MS/MS data was analyzed using Proteome Discoverer 1.4.1.14 (Thermo Fisher Scientific). For database searching, raw MS files were submitted to the *Homo sapiens* UniProtKB/Swiss-Prot database (HUMAN_Uniprot_R_2017_06.fasta; released June 2017, 23,064 protein entries) with *Sus scrofa* Trypsin added to it using SEQUEST HT version 28.0 (Thermo Fisher Scientific). For re-scoring, percolator search node was used. Searches were performed using the following parameters: two maximum missed cleavage sites for trypsin, TMT-labeled N-terminus and aminoacids (lysine, histidine, serine, and threonine) (+229.163 Da) and methionine oxidation (+15.995 Da) as dynamic modifications, cysteine carbamidomethylation (+57.021 Da) as static modification, 20 parts per million (ppm) precursor mass tolerance, 0.1 Da fragment mass tolerance, 10 milli mass units (mmu) peak integration tolerance, and most confident centroid peak integration method. Percolator was used for protein identification with the following identification

criteria: at least one unique peptide per protein with a FDR of <1 %. Different isoforms of the same protein were treated as ungrouped to avoid any possible ambiguity (Amaral et al. 2014, Bogle et al. 2017, Barrachina et al. 2019).

For protein quantification data, normalized TMT quantitative values were obtained from each identified spectrum derived from the ratio of the intensity of reporter ions from HCD-MS2 spectra corresponding to each individual sample (TMT-127 to TMT-131) with the internal control (TMT-126), obtained by Proteome Discoverer software. In order to obtain reliable results, only those proteins with at least 1 unique peptide identified with FDR <1 %, quantified by $\geqslant$ 2 peptide spectrum matches (PSM) in all the samples, and with a coefficient of variation <50 % in at least 75 % of the samples were considered for further statistical analyses.

Western blotting

Western Blot (WB) and relative semiquantification of protein abundance by comparative densitometry were carried out as previously described (Ferrer-Vaquer et al. 2016). Denatured protein extracts corresponding to 1 million sperm were used. Two different proteins were selected as normalizers, α-tubulin (TUBA) as cytoplasmic and Succinate Dehydrogenase Complex Flavoprotein Subunit A (SDHA; mitochondrial) as mitochondrial normalizers, using specific antibodies at indicated dilutions: mouse monoclonal anti-TUBA antibody (1:5,000; clone DM1A; T6199, Sigma-Aldrich, St. Louis, MO, USA) and rabbit polyclonal anti-SDHA antibody (1:5,000; ab14715, Abcam, Cambridge, MA, USA).

To validate proteomics results in terms of relative abundance, we used specific antibodies raised against Proteasome Subunit Alpha 1 (PSMA1) (rabbit polyclonal; 1:2,000; HPA037646, Sigma-Aldrich) and dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex (DLAT) (mouse monoclonal; 1:1,000; ab110333, Abcam), which were normalized by re-blotting against TUBA (mouse monoclonal; 1:5,000; T6199, Sigma) and SDHA (mouse monoclonal; 1:5,000; ab14715, Abcam). Immunoblot detection was performed with peroxidase labeled anti-mouse (NIF825, GE Healthcare Bio-Sciences, Marlborough, MA, USA) or anti-rabbit (NIF824, GE Healthcare Bio-Sciences) as secondary antibodies at 1:10,000. Immunoreactivity was detected using chemiluminescent substrate (Luminata Classico, Millipore) and exposure to X-ray films. Densitometric quantification was performed using a specific plug‐in of ImageJ software (National Institutes of Health, Bethesda, MD).

Determination of mitochondrial DNA copies per genome in sperm

Total DNA, including genomic (gDNA) and mitochondrial (mtDNA) was extracted from 2 million sperm after 50 % DGC processing and quantified as previously described (Torra-Massana et al. 2018). The relative abundance of mtDNA copies per genome was determined by real-time quantitative PCR (qPCR). Two different mitochondrial genes, Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 1 (*MT-ND1*) and 4 (*MT-ND4*), were assessed and quantified against the nuclear gene Beta-2 Microglobulin (*B2M*). Primer sequences are listed in Table S3. Each qPCR reaction included 1 ng of total DNA, 10 µl of 2× SsoAdvanced Universal SYBR Green Supermix (BioRad, Hercules, CA, USA), and 0.1 µl of each primer (from a 100 µM stock), to a final reaction volume of 20 µl. The qPCR program used consisted of an initial denaturation step of 30 s at 95 °C and 40 cycles of 95 °C for 5 s and 60 °C for 30 s. All qPCRs were carried out in triplicate using CFX96 (BioRad). The mitochondrial genes *MT-ND1* and *MT-ND4* were normalized against *B2M* using the ΔΔCq method. Primer pair efficiencies were considered for the analysis, which were evaluated using sequential dilutions of human total DNA from sperm.

Analysis of mitochondrial membrane potential by flow cytometry

Sperm mitochondrial membrane potential (MMP) was measured by using JC-1 (Thermo Fisher Scientific), as previously described (Fernandez and O'Flaherty 2018). Briefly, 2 million sperm purified by 50 % DGC were stained with 2 µM JC-1 in PBS for 15 minutes in the dark at 37°C. Immediately before applying the sample to flow cytometry evaluation, SYTOXTM Blue (Thermo Fisher Scientific) was added at 1 mM. All samples were analyzed using a Sony SA3800 flow cytometer (Sony Biotechnology, San Jose, CA, USA), and data analyzed using FlowJo software v. 10.5.2 (LLC, Ashland, OR, USA). SYTOXTM Blue staining was used to distinguish the live sperm population and, within this population, JC-1 staining identified the sperm population presenting MMP (green+/red+ population). For each sample, the MMP was expressed as the ratio between the mean intensity of red fluorescence and green fluorescence within live sperm population, expressed as relative units (RU).

Determination of sperm proteasome activity

Proteasome activity in sperm samples was determined using the Proteasome Activity Assay Kit (ab107921, Abcam) following the manufacturer's instructions. Briefly, for each individual sample, 2 million sperm after 50% DGC were lysed in 20 µl of 1 % NP-40 by gentle pipetting and centrifuged for 15 minutes (4ºC) at 15,800 g, to recover the detergent-soluble native protein complexes (supernatant). A standard curve was prepared using

AMC standard provided by the manufacturer, and all samples were assayed in duplicate, with and without proteasome inhibitor (MG-132, Abcam), in a 96-well black-opaque plate. The proteasome activity kinetics was determined by incubating the extracts with an AMC-tagged peptide at 37 °C during 70 minutes in a FLx800™ microplate reader (BioTek, Winooski, VT, USA). The AMC release was monitored by measuring AMC fluorescence every 10 minutes (λex = 360 nm; λem = 460 nm). Proteasome activity is expressed as proteasome activity units (PAU), where one unit is defined as the amount of extract from 10^6 sperm which generates 1 nmol of AMC per minute at 37 °C.

Statistical analyses

Statistical analyses were conducted using R software version 3.4.4 (http://www.r-project.org) and statistical package IBM SPSS 18.0 (New York, USA). Statistical significance was set at $p < 0.05$.

The sperm proteomic differences between controls and FF were determined by group analysis, using a two tailed Student's t-test ($p<0.05$) on normalized (log2 transformation) relative quantification proteomic data at both the protein and the peptide levels. After this analysis, a post-hoc power analysis was conducted using G*Power 3 software (Faul, F., et al. 2007) to assess the sample-size for validation. Additionally, FF-specific altered proteins were determined in individual FF samples, by two different statistical approaches. The first approach was based on identifying alterations of the stable-correlated proteins established in the control population in individual patients (FF) as previously described (Barrachina et al. 2019). Briefly, the intensity values from HCD-MS2 spectra corresponding to each control sample were used first to establish the stable-protein pair pattern for control group. Specifically, stable-protein pairs were determined by applying the statistical principle of 2 proteins (with more than 1 peptide quantified under strict criteria for each one; n=113 proteins) were highly correlated when $\geqslant$ 75 % of the possible peptide combinations had a Pearson correlation coefficient $\geqslant$ 0.9 (Fig. S1). Afterwards, in order to determine alterations in individual samples, stable-protein pairs analysis for the control group was repeated by adding one FF patient once at a time. Only those proteins which lost >50 % of correlations previously determined in the control group were considered as altered in a particular FF sample. The second approach was based on the identification of proteins with an extreme high or low abundance for specific samples through the application of high strict ROUT's test (0.2 % threshold).

Student's t-test was used to compare the mean of Western Blot protein semiquantification, MMP, proteasome activity and mtDNA copies per genome between controls and FF samples. Statistical significance was set at $p < 0.05$.

RESULTS

Semen parameters and clinical outcomes

Individual demographic and seminal parameters, as well as fertilization outcomes for all patients included in the study are indicated in Table S1. The unique significant difference between both groups was the mean fertilization rate (83.19 % in controls vs. 3.05 % in FF) (Table 1). All patients were normozoospermic, except for F4 and C4, which presented mild asthenozoospermia and teratozoospermia, respectively (Table S1). All samples presented <1 % contaminating cells after 50 % DGC processing (Table 1).

General sperm proteome analysis

By using TMT labelling and high-resolution 2D-LC-MS/MS analysis, a total of 6,662 unique peptides with <1 % FDR were identified, comprising 1,394 different proteins. As an example, the spectrum for peptide GIDLTQVK of protein DLAT is shown in Fig. S2. A total of 231 sperm proteins were quantified in high confidence according to our previously established strict quantification criteria and only those ones were considered for subsequent statistical analyses (Table S4).

Group analysis identifies differentially abundant proteins related to mitochondrial and proteasomal function

In order to identify potential biomarkers for FF, a comparison of relative protein quantification values between FF and control groups was performed. This analysis identified 9 proteins presenting changes in their abundance between FF and control groups (t-test, $p<0.05$) (Table 2). Four of these proteins presented lower levels in the FF group: FMR1 Neighbor (FMR1NB), Family with Sequence Similarity 209 Member B (FAM209B), Ras-related protein Rab-2b (RAB2B) and Proteasome subunit alpha 1 (PSMA1, a protein essential for proteasome function). Additionally, five mitochondrial proteins were found more abundant in FF samples: Dihydrolipoyllysine-residue acetyltransferase component of PDH (DLAT), ATP Synthase Peripheral Stalk Subunit D (ATP5H), Solute Carrier Family 25 Member 3 (SLC25A3), Solute Carrier Family 25 Member 6 (SLC25A6), and Fumarate Hydratase (FH).

Of the 9 proteins detected with an altered abundance in FF samples, 3 of them were reliable quantified with more than 1 peptide (DLAT, FAM209B and FH). Whereas just

Table 1. Demographical and seminal parameters, clinical variables and density gradient processing information for the two groups included in the study (FF and controls). Results are indicated as mean ± SD [range]. Chi-squared test was used to compare the fertilization rates between control and FF groups, while Student's t-test was used for the rest of comparisons.

	FF group (n = 5)	Control group (n = 12)	p-value
Male age; years	40.0 ± 3.7 [36.0 - 43.0]	37.3 ± 3.8 [31.0 - 43.0]	0.1989
Male BMI, kg/m^2	23.3 ± 1.1 [22.4 - 24.5]	24.2 ± 2.6 [21.3 - 29.0]	0.4732
Ejaculate volume, ml	5.3 ± 0.9 [4.1 - 6.4]	4.6 ± 1.4 [1.6 - 6.7]	0.3226
Concentration, million/ml	86.5 ± 70.3 [21.8 - 191.0]	118.6 ± 64.6 [18.5 - 242.2]	0.3765
Sperm motility, % a+b	54.7 ± 28.7 [21.3 - 87.1]	62.0 ± 16,6 [34.4 - 87.2]	0.5144
Sperm morphology, % normal	7.8 ± 2.9 [4.0 – 12.0]	5.7 ± 2.8 [2.0 – 13.0]	0.1832
Inseminated oocytes per cycle, n	6.0 ± 3.4 [3.0 - 10.8]	7.2 ± 3.3 [3.0 - 16.0]	0.5083
Age of woman providing the oocytes*, years	34.9 ± 4.9 [27.0 – 42.0]	30.0 ± 5.0 [20.0 – 40.0]	0.0840
FR, % (2PN / injected oocytes)	3.1 (4 / 131)	83.2 (99 / 119)	< 0.0001
Contaminating cells after DGC, %	0.21 ± 0.16 [0.00 - 0.37]	0.23 ± 0.16 [0.0 - 0.52]	0.8175

*Oocyte age is considered the age of the female partner, or the age of the oocyte donor when using donor oocytes. BMI, Body Mass Index; DGC, Density gradient centrifugation; FR, Fertilization rate; PN, pronuclei; SD, Standard deviation

one of the peptides quantified presented differential abundance for FAM209B (1 out of 2 peptides) and FH (1 out of 4 peptides), all the peptides of DLAT (2 out of 2 peptides) showed significant differential abundance between controls and FF ($p < 0.05$), (Fig. S3).

Analysis of mitochondrial protein abundance and membrane potential in FF samples

The significant increase in the abundance of five mitochondrial proteins in FF patients was initially validated by WB against DLAT (which showed significant differential abundance for the 2 identified peptides) using TUBA as normalizer (The sample-size calculation for a power >80% predicted the inclusion of a minimum of 3 and 7 samples for FF and control groups, respectively). Additionally, in order to evaluate whether observed differences in mitochondrial proteins abundance between samples could be a general reflect of the number of mitochondria, the abundance of DLAT was assessed relative to a specific mitochondrial protein, the mitochondrial protein SDHA (Fig. 2A and Fig. S4). The highest abundance of DLAT was observed independently of the use of either general or mitochondrial specific normalizer. Complementary, in order to assess whether high protein abundance in FF samples associates to alterations in mtDNA content, the mean value of mtDNA per sperm genome (gDNA) was determined by qPCR. However, no significant differences were observed between controls and FF samples (Fig. 2B), suggesting that the higher abundance of specific mitochondrial proteins observed in FF samples was not associated with mtDNA content.

As the differentially abundant mitochondrial proteins detected participate in the mitochondrial tricarboxylic acid cycle (TCA cycle), mitochondrial membrane potential (MMP) in sperm was further assessed by flow cytometry to gain insight into mitochondrial dysregulations in FF samples. We specifically assessed MMP on live sperm population (SYTOX^{TM} Blue negative population). In FF patients there was a lower percentage of sperm presenting MMP (59.1 ± 12.0 %) compared to controls (72.0 ± 20.3 %), but this difference did not reach statistical significance ($p = 0.274$) (Fig. 2C). However, the mean values of MMP within this population, (SYTOX -, red+/green+) did not present significant differences between controls and FF: 132.5 ± 47.9 RU vs. 115.8 ± 9.5 RU, respectively (Fig. 2D).

Table 2. Identified proteins showing significant differential abundance between samples, and the corresponding post-hoc power analysis, with repetitive fertilization failure after ICSI and controls (>75% mean fertilization rate) at group level, following strict criteria analysis.

UniProt accession number	Protein Symbol	Protein name	p-value*	Relative abundance (FF : control)	Post-hoc Power
P10515	DLAT	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	0.0033	1.257	0.94
Q00325	SLC25A3	Phosphate carrier protein, mitochondrial	0.0075	1.181	
Q8N0W7	FMR1NB	Fragile X mental retardation 1 neighbor protein	0.0123	0.868	
Q8WUD1	RAB2B	Ras-related protein Rab-2B	0.0187	0.866	
Q5JX69	FAM209B	Protein FAM209B	0.0226	0.872	
P25786	PSMA1	Proteasome subunit alpha type-1	0.0341	0.893	0.45
O75947	ATP5H	ATP synthase subunit d, mitochondrial	0.0407	1.370	
P12236	SLC25A6	ADP/ATP translocase 3, mithocondrial	0.0433	1.128	
P07954	FH	Fumarate hydratase, mitochondrial	0.0473	1.098	

*Student t-test.

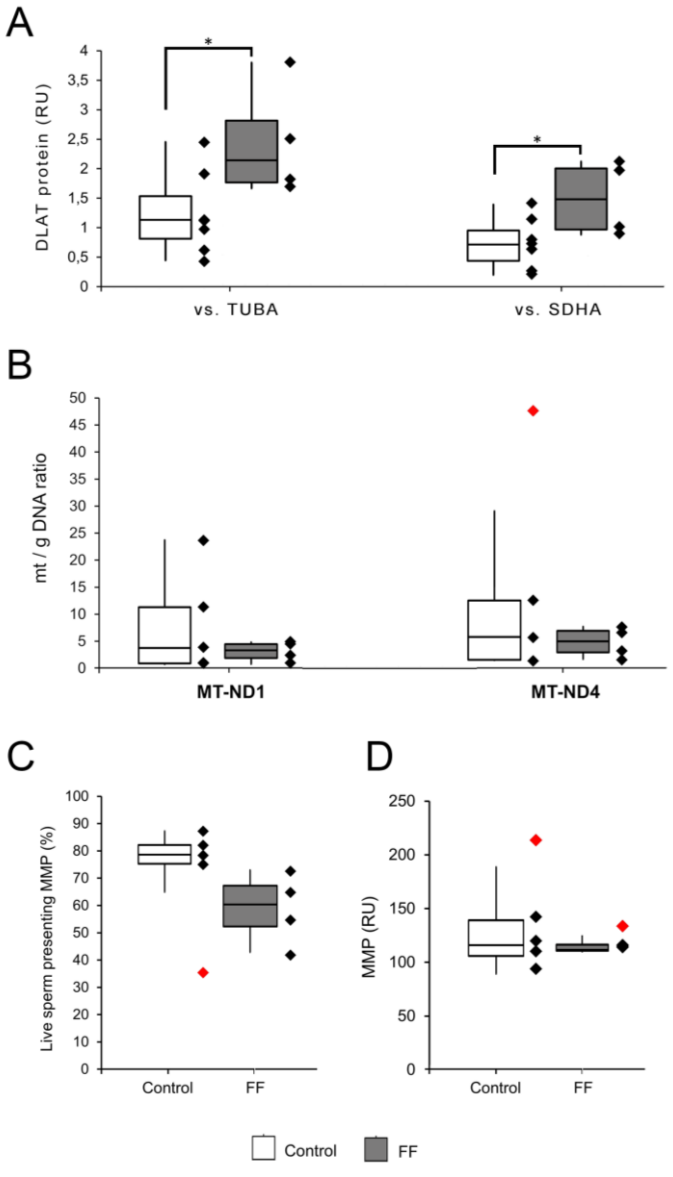


Figure 2. Analysis of mitochondrial protein abundance and function in controls and FF human sperm samples. Red dots represent outliers. **A**. Mitochondrial protein abundance was analyzed by semiquantitative Western Blot. The differential abundance observed for DLAT protein in controls (n = 7; C1 and C3-C8) and FF samples (n = 4; F1-F4) was validated using two independent normalizers, the mitochondrial protein SDHA and the cytoplasmic protein TUBA. Values are expressed as the ratio of semiquantitative densitometric units measured using ImageJ. Asterisk indicates significant difference (t-test, $p < 0.05$). **B**. Relative mtDNA abundance of controls (n = 5; C7 and C9-C12) and FF samples (n = 4; F1 and F3-F5). Two mitochondrial genes (MT-ND1 and MT-ND4) were normalized against a nuclear gene (B2M) using the ΔΔCq method. **C**. Box plot representing the percentage of live sperm by SYTOX™ Blue assay SYTOX (SYTOX negative). **D**. Box plot representing the mean MMP values, expressed as the ratio red / green JC-1 fluorescence within live sperm population; relative units, RU) in FF (n = 4; F1 and F3-F5)) and control (n = 5; C7 and C9-C12) samples.

Analysis of proteasome protein abundance and activity in FF samples

in accordance with the proteomics results, significant lower levels of PSMA1 were detected in FF samples by WB (Fig. 3A and Fig. S4), although due to sample limitation, not enough samples could be assessed to reach a power > 80 %. Since PSMA1 belongs to the 20S proteasomal core, we tested whether the low abundance of PSMA1 in FF samples could affect proteasomal activity in sperm from controls and FF samples by means of a fluorometric assay. Proteasome activity in sperm was depleted when incubated in presence of MG132 inhibitor (Fig. S5), indicating that the measure performed corresponds to specific proteasome activity in sperm. Although sperm samples showed high variability in the proteasome activity levels, no significant differences were observed between controls and FF samples (0.633 ± 0.264 and 0.382 ± 0.136 PAU, respectively, p = 0.131) (Fig. 3B).

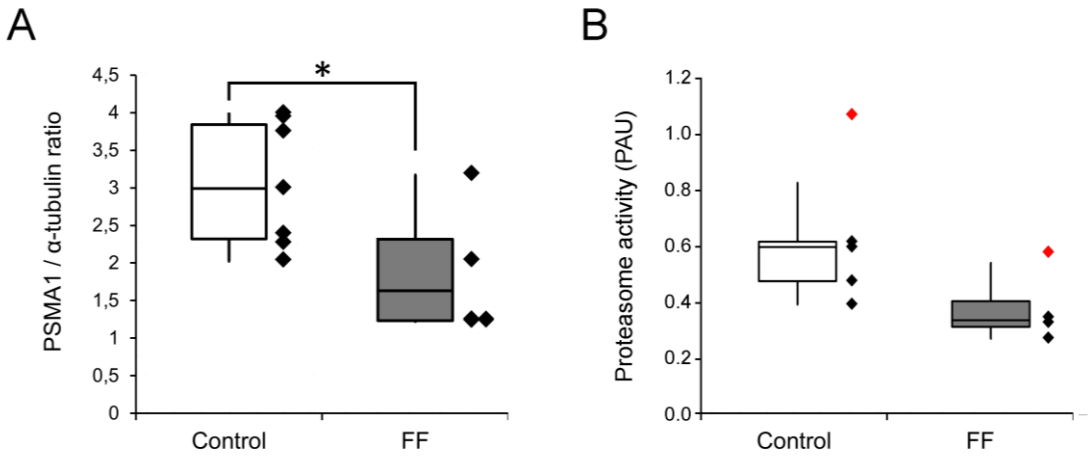


Figure 3. Analysis of proteasomal protein levels and proteasome activity in sperm. Red dots represent outliers. **A**. Validation by Western Blot of the differential abundance observed for PSMA1 in controls (n = 7; C1 and C3-C8) and FF samples (n = 4; F1-F4), using TUBA as normalizer. Semiquantitative ratios are expressed as relative units. Asterisk indicates significant difference (t-test, $p < 0.05$). **B**. Box plot showing proteasome activity levels for samples with good fertilization rates (controls, n = 5; C7 and C9-C12) and fertilization failure (FF, n = 4; F1 and F3-F5) after ICSI. Results are expressed as proteasome activity units (PAU), which are defined as the amount of extract in 1,000,000 sperm which generates 1 nmol of AMC per minute at 37°C.

Stable-protein pairs and outlier's profile analysis identifies specific alterations in each individual sample

Although group analysis pointed out that sperm mitochondria were relevant for oocyte fertilization, further insights were obtained with the analysis of individual FF samples by the establishment of stable-protein pairs and outliers on proteomics results. A total of 73 stable-protein pairs involving 28 different proteins (related to sperm structure, metabolism and fertilization) were identified for the control group with high fertilization rates (Table S5). The reassessing of these stable-protein pairs in control group but adding one FF sample once a time revealed that F1 maintained all the stable-protein pairs established previously in the control population. In contrast, several stable-protein pairs were lost for the other FF patients (F2, F3 and F4) (Table 3). Interestingly, all

Table 3. Altered stable-protein correlations in individual FF patients compared to control population (≥ 75% fertilization rate after ICSI; n=8).

Sample	Number of correlations	Proteins involved	Proteins that lost ≥ 50 % correlations
F1	93	32	-
F2	47	19	PDHA2, ACRV1, PHB2, RUVBL1, ZPBP, ASRGL1, NUP210L, FAM209B, DECR1, ATP5F1D, PDHB, TEKT2
F3	57	24	RPLP2, PHB2, FAM209B, ATP5F1D, TEKT2
F4	42	21	CD59, OXCT1, NUP54, CCT3, ACRV1, SPACA4, RPLP2, PHB2, RUVBL1, MPC1L, NUP210L, FAM209B, DECR1, ATP5F1D, PDHB, TEKT2

correlations established in the control population for another mitochondrial protein related to energy production, the ATP Synthase F1 Subunit Delta (ATP5F1D), a member of the ATP synthase family as ATP5H, were lost in F2, F3 and F4 samples (Table 3).

The outliers' analysis also provided valuable information about proteins with extremely altered levels in individual FF samples, indicating the high molecular heterogeneity leading to a similar phenotype (Table S6). Whereas none of the control samples displayed an extremely altered abundance in any of the proteins identified and

quantified, 5 outliers were identified in 2 of the 4 FF samples (F1 and F4). For instance, two proteins involved in the mitochondrial electron transport chain, cytochrome c oxidase subunit 5B (COX5B) and cytochrome b-c1 complex subunit 2 (UQCRC2), showed extreme high levels in samples F1 and F4, respectively. Additionally, F1 also displayed extreme high levels of the protein sperm acrosome-associated protein 7 (SPACA7).

DISCUSSION

To the best of our knowledge, this is the first study using comparative proteomics in individual sperm samples to identify markers for fertilization failure after ICSI due to a male factor. Our study, which combines conventional (group) and novel (individualized) approaches to analyze quantitative proteomic data, such as the recent established stable-protein pair analysis (Barrachina et al. 2019), aims to contribute to the elucidation of the molecular and cellular mechanisms behind FF. The combination of both grouped and individualized statistical analysis identified certain alterations in protein levels in FF samples related to sperm metabolism and spermatogenesis.

Mitochondria participate in various cellular functions, including ATP production, calcium homoeostasis and generation of reactive oxygen species (ROS), which are crucial for mammalian sperm. However, its exact role in the mammalian fertilization has not been elucidated yet. It is well-known that the oxidative respiration taking place in mitochondria is essential for sperm motility and capacitation (Rogers et al. 1977, Ramio-Lluch et al. 2014), and functional mitochondrial parameters have been clearly correlated with human sperm fertilization ability (reviewed in (Amaral et al. 2013). Moreover, alterations in mitochondrial function have also been associated with higher sperm apoptosis rates, impaired sperm calcium signalling, and higher sperm DNA damage, therefore, compromising the formation of male pronucleus after the spermatozoon enters into the oocyte cytoplasm (Reyes et al. 1993). Herein, we show for the first time that sperm from male patients presenting FF after ICSI could have alterations in the levels of specific mitochondrial proteins.

Mitochondrial proteins, such as DLAT, ATP5H, SLC25A3, SLC25A6 and FH, displayed significant higher amount levels in the FF group. Additionally, other mitochondrial proteins such as ATP5F1D, PDHA2, PDHB, PHB2 or OXCT1 appeared in altered abundance in 3 out the 4 FF patients assessed by after the stable-protein pair analysis approach. All these proteins are directly linked to the oxidative phosphorylation network (OxPHOS). For instance, among these altered mitochondrial proteins, PDHA2, PDHB and DLAT are part of the pyruvate dehydrogenase complex (PDC), a component of the pyruvate and citric acid cycle. PDC is organized in three main complexes: (i) pyruvate dehydrogenase (E1 subunit: PDHA1, PDHA2, PDHB), (ii) dihydrolipoyltransacetylase

(E2 subunit: DLAT), and (iii) dihydrolipoamide dehydrogenase (E3 subunit: DLD). One of the main causes of mitochondrial dysfunction could rely on defects on the regulation of the mitochondrial pyruvate metabolism and PDC, which acts as modulator of metabolic homeostasis (Gray et al. 2014, Jeoung et al. 2014). Moreover, the disruption of PDC function in hamster sperm results in fertilization failure due to the inability to elicit calcium oscillation in the oocyte (Siva et al. 2014). Additionally, FH is involved in NADH and FADH2 production (by TCA cycle), while ATP5H and ATP5F1D belong to Complex V of the respiratory chain to produce ATP and SLC25A3 and SLC25A6 to Pi and ATP/ADP transport, respectively. All these processes lead to the establishment of a mitochondrial membrane potential (MMP) needed to maintain the correct mitochondrial function. While previous studies showed association between conventional sperm parameters and MMP values (Sousa et al. 2011, Zhang et al. 2016), no correlation with fertilization rates have already been described. As it has been described in mammalian sperm that freeze-thaw procedures can affect viability and MMP values (Castro et al. 2016), we decided to analyze live sperm in both, controls and FF patients, after thawing procedures. Altogether, our results suggest a slight disturbance on the percentage of live sperm presenting MMP in FF samples, although no statistical differences were observed.

Different studies have revealed that mitochondria and proteasome are closely related by maintaining a crosstalk: abnormal mitochondrial function may be linked to a reduced proteasome system activity and vice versa, causing cellular dysfunction, cellular stress and impaired cellular homeostasis and viability (Maharjan et al. 2014, Segref et al. 2014, Ross et al. 2015, D'Amico et al. 2017). Proteasome dependent proteolysis plays key roles not only during spermatogenesis, but also in sperm capacitation, sperm intracellular calcium signaling, acrosome reaction and oocyte fertilization in mammals (Sutovsky 2011, Kerns et al. 2016, Zigo et al. 2018). Inside the oocyte, proteasomal function is required for the degradation of mitochondrial piece and tail connecting piece, the release of sperm centrosome, the proper aster formation and pronuclear development in both bovine and human models. In addition, half of the oocytes that fail to be fertilized after ICSI seem to present an abnormal accumulation of proteasomes near the non-decondensed male nucleus, which also fail to develop PN (Rawe et al. 2008). Interestingly, PSMA1, a component of the proteasome core, showed lower levels in the FF patients. Although no significant differences in sperm proteasomal activity between controls and FF samples were detected, post-hoc power and sample-size calculations indicated that further research is needed to confirm whether sperm proteasomal activity is a factor that affects the fertilization ability after ICSI.

Three additional proteins (FAM209B, FMR1NB and RAB2B) were also found reduced in abundance in FF patients, the first one also confirmed as dysregulated in 3 FF patients

through the stable-protein pair analysis. Two of them, FAM209B and FMR1NB are highly expressed in testicular tissue according to The Human Protein Atlas database (Uhlen et al. 2015), but their cellular function has not been described. FMR1NB-KO mice do not show fertility problems (Miyata et al. 2016). However, we cannot discard that, in humans, alterations in their expression levels during spermatogenesis could be a cause of FF. Moreover, RAB2B is a member of the GTPase family, which participates in the regulation of Golgi morphology in human cells (Aizawa and Fukuda 2015). Therefore, as the acrosome is an exocytotic organelle derived from the Golgi apparatus, RAB2B deficiency may lead to abnormal acrosomal formation, as well as some structural defects, which could compromise sperm fertilization ability after ICSI.

In addition to group analysis, individual analyses based on the stable-protein pairs and outlier's profiles have provided valuable information about each specific FF patient. For instance, although patient F1 maintained all the stable correlations found in patients with high fertilization rate, suggesting a similar general proteomic signature, the outliers' analysis revealed three proteins with extreme higher levels in this specific patient. First, mitochondrial COX5B (belonging to Complex IV), corroborating our hypothesis about the importance of mitochondrial respiration for a successful fertilization even after ICSI. Second, SPACA7, an acrosomal protein previously related to fertilization in mouse and with unknown functions after sperm penetration (Nguyen et al. 2014). Third, HSPE1, a mitochondrial chaperonin with important roles during capacitation in mouse sperm (Walsh et al. 2008).

For the other 3 FF samples (F2, F3 and F4), in contrast, stable-protein pairs analysis uncovered dysregulated proteins which were stable within the control group analysis, such as the aforementioned FAM209B and components of the PDC and ATP synthase. Stable-protein pairs analysis also showed an impairment in the levels of PHB2, which is essential for sperm mitochondria ubiquitination and mitophagy inside the oocyte (Sutovsky et al. 1999); TEKT2 the expression levels of which were positively associated with fertilization rates after ICSI (Bhilawadikar et al. 2013); and NUP54 and NUP210L, two proteins associated with the nuclear envelope which are required for pronucleus formation and interaction with sperm aster during oocyte fertilization (Payne et al. 2003).

We recognize that our study has some limitations. First, due to cases of repetitive fertilization failure after ICSI that accomplished the inclusion and exclusion criteria indicated in the study design are very rare, our sample size was limited to 5 idiopathic FF samples. We performed post-hoc power and sample-size calculations based on proteomic results and realize that additional samples would be beneficial to be included to validate some of the putative markers. However, similar sample sizes have been employed by others to identify specific molecular causes of infertility by high-throughput

proteomics (Frapsauce et al. 2014, Guo et al. 2019, Murdica et al. 2019). Second, since during sperm capacitation several proteins, and specifically mitochondrial proteins, are reported to undergo post-translational modifications (PTMs) (Shivaji et al. 2009), complementary studies targeting protein PTMs (Castillo et al. 2019), as well as other modifications such as gene sequence variants, or alternative splicing, will be useful to understand all the mechanisms potentially involved in sperm-dependent ICSI fertilization failure. Finally, the female factor cannot be ruled out in FF after ICSI. Some oocyte factors, such as nucleus-cytoplasmic maturation asynchrony and spindle defects may explain some cases of FF after ICSI (Swain and Pool 2008, Yeste et al. 2016).

In conclusion, our study reveals additional mechanisms potentially involved in male-related fertilization failure after ICSI, and suggests the process is indeed multifactorial. In addition, we reinforce the idea that molecular and cellular events occurring in sperm before and during oocyte fertilization, such as sperm capacitation or mitochondrial metabolism, could have profound effects not only on sperm ability to reach and penetrate the oocyte, but also on oocyte activation and early embryo development. Likewise, specific alterations in mitochondrial proteins could be the underlying cause in some FF cases. Further studies are also required to define the value of the suggested proteins as diagnostic and prognostic biomarkers for fertilization failure after ICSI.

DATA AVAILABILITY

Supplementary Data has been deposited to OneDrive repository from the University of Barcelona. The data files can be downloaded at the following link until published: https://ubarcelona-my.sharepoint.com/:f:/r/personal/ada_soler_ub_edu/Documents/Annex%203?csf=1&web=1&e=9G99Yu

ACKNOWLEDGEMENTS

We would like to thank the laboratory staff from Clínica Eugin for their help in sample handling, Désirée García for statistical support, Filippo Zambelli, Judit Castillo and Ferran Barrachina for experimental counseling, and Gustavo Tiscornia for critical revision of the manuscript. This work was supported by intramural funding of Clinica EUGIN, and by the Secretary for Universities and Research of the Ministry of Economy and Knowledge of the Government of Catalonia (GENCAT 2015 DI 049) to M. T-M. EUGIN-UB Research Excellence Program (EU -REP 2014), the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad; Fondos FEDER 'Una manera de hacer Europa', PI 13 /00699, and PI 16 /00346), from Fundación Salud 2000 (SERONO 13 –015) to R.O. M.J. is supported by the Government of Catalonia (Generalitat de Catalunya, Pla Estratègic de Recerca i Innovació en Salut, PERIS 2016 -2020, SLT SLT002/16 /00337).

DISCLOSURES

No competing interest declared.

AUTHORS' CONTRIBUTION

M.T-M: data collection and analysis, results interpretation, critical discussion and manuscript writing. M.J.: data analysis, results interpretation, critical discussion and manuscript writing. M.B.: project conception, study design, data collection and analysis, results interpretation, critical discussion and manuscript writing. A. S-V.: data collection and analysis and manuscript editing. D. D-D.: data analysis and interpretation, manuscript revision. A.R: project conception, critical discussion of the manuscript. R.O: project conception, study design and critical discussion. Also provided critical review and editing of the manuscript. R.V: project conception, study design, and critical discussion. Also provided critical review and editing of the manuscript.

REFERENCES

Aizawa, M., Fukuda, M., 2015. Small gtpase rab2b and its specific binding protein golgi-associated rab2b interactor-like 4 (gari-l4) regulate golgi morphology. The Journal of biological chemistry. 290, 22250-61

Amaral, A., et al., 2013. Mitochondria functionality and sperm quality. Reproduction. 146, R163-74

Amaral, A., et al., 2014. Identification of proteins involved in human sperm motility using high-throughput differential proteomics. Journal of proteome research. 13, 5670-84

Barrachina, F., et al., 2019. Stable-protein pair analysis as a novel strategy to identify proteomic signatures: Application to seminal plasma from infertile patients. Molecular & cellular proteomics : MCP. 18, S77-S90

Bhilawadikar, R., et al., 2013. Levels of tektin 2 and catsper 2 in normozoospermic and oligoasthenozoospermic men and its association with motility, fertilization rate, embryo quality and pregnancy rate. Journal of assisted reproduction and genetics. 30, 513-23

Bogle, O.A., et al., 2017. Identification of protein changes in human spermatozoa throughout the cryopreservation process. Andrology. 5, 10-22

Boulet, S.L., et al., 2015. Trends in use of and reproductive outcomes associated with intracytoplasmic sperm injection. Jama. 313, 255-63

Castillo, J., et al., 2019. Proteomic changes in human sperm during sequential in vitro capacitation and acrosome reaction. Frontiers in cell and developmental biology. 7, 295

Castillo, J., et al., 2018. The contribution of human sperm proteins to the development and epigenome of the preimplantation embryo. Human reproduction update. 24, 535-555

Castro, L.S., et al., 2016. Sperm cryodamage occurs after rapid freezing phase: flow cytometry approach and antioxidant enzymes activity at different stages of cryopreservation. Journal of Animal Science and Biotechnology. 7, 17

D'amico, D., et al., 2017. Cytosolic proteostasis networks of the mitochondrial stress response. Trends in biochemical sciences. 42, 712-725

Fariello, R.M., et al., 2012. Effect of smoking on the functional aspects of sperm and seminal plasma protein profiles in patients with varicocele. Human reproduction. 27, 3140-9

Faul, F., et al., 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175-191

Fernandez, M.C., O'flaherty, C., 2018. Peroxiredoxin 6 is the primary antioxidant enzyme for the maintenance of viability and DNA integrity in human spermatozoa. Human reproduction. 1;33(8):1394-1407

Ferrer-Vaquer, A., et al., 2016. Plczeta sequence, protein levels, and distribution in human sperm do not correlate with semen characteristics and fertilization rates after icsi. Journal of assisted reproduction and genetics. 33, 747-56

Flaherty, S.P., et al., 1998. Fertilization failures and abnormal fertilization after intracytoplasmic sperm injection. Human reproduction. 13 Suppl 1, 155-64

Frapsauce, C., et al., 2014. Proteomic identification of target proteins in normal but nonfertilizing sperm. Fertility and sterility. 102, 372-80

Gray, L.R., et al., 2014. Regulation of pyruvate metabolism and human disease. Cellular and molecular life sciences : CMLS. 71, 2577-604

Guo, Y., et al., 2019. Proteomic analysis of dpy19l2-deficient human globozoospermia reveals multiple molecular defects. Proteomics. Clinical applications. 13, e1900007

Intasqui, P., et al., 2013. Unraveling the sperm proteome and post-genomic pathways associated with sperm nuclear DNA fragmentation. Journal of assisted reproduction and genetics. 30, 1187-202

Jeoung, N.H., et al., 2014. Regulation of pyruvate metabolism in metabolic-related diseases. Reviews in endocrine & metabolic disorders. 15, 99-110

Jodar, M.B., F.; Oliva, R., 2017. The use of sperm proteomics in the assisted reproduction laboratory. In: Garrido n, rivera r, editors. Practical guide to sperm analysis. Boca raton (fl): Crc press; p. 233–244.

Kashir, J., et al., 2018. Phospholipase c zeta and calcium oscillations at fertilisation: The evidence, applications, and further questions. Advances in biological regulation. 67, 148-162

Kerns, K., et al., 2016. Regulation of sperm capacitation by the 26s proteasome: An emerging new paradigm in spermatology. Biology of reproduction. 94, 117

Kupka, M.S., et al., 2014. Assisted reproductive technology in europe, 2010: Results generated from european registers by eshredagger. Human reproduction. 29, 2099-113

Légaré, C.D., A.; Fournier, F.; Bourassa, S.; Force, A.; Cloutier, F.; Tremblay, R.; Sullivan R., 2014. Investigation of male infertility using quantitative comparative proteomics. . Journal of proteome research. 13:5403-14

Liao, T.T., et al., 2009. Proteome analysis of round-headed and normal spermatozoa by 2-d fluorescence difference gel electrophoresis and mass spectrometry. Asian journal of andrology. 11, 683-93

Liu, X., et al., 2018. Itraq-based analysis of sperm proteome from normozoospermic men achieving the rescue-icsi pregnancy after the ivf failure. Clinical proteomics. 15, 27

Maharjan, S., et al., 2014. Mitochondrial impairment triggers cytosolic oxidative stress and cell death following proteasome inhibition. Scientific reports. 4, 5896

Miyata, H., et al., 2016. Genome engineering uncovers 54 evolutionarily conserved and testis-enriched genes that are not required for male fertility in mice. Proceedings of the National Academy of Sciences of the United States of America. 113, 7704-10

Murdica, V., et al., 2019. Proteomic analysis reveals the negative modulator of sperm function glycodelin as over-represented in semen exosomes isolated from asthenozoospermic patients. Human reproduction. 34, 1416-1427

Nasr-Esfahani, M.H., et al., 2007. Effects of failed oocyte activation and sperm protamine deficiency on fertilization post-icsi. Reproductive biomedicine online. 14, 422-9

Nguyen, E.B., et al., 2014. Spaca7 is a novel male germ cell-specific protein localized to the sperm acrosome that is involved in fertilization in mice. Biology of reproduction. 90, 16

Nozawa, K., et al., 2018. Sperm-borne phospholipase c zeta-1 ensures monospermic fertilization in mice. Scientific reports. 8, 1315

Palermo, G., et al., 1992. Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte. Lancet. 340, 17-8

Payne, C., et al., 2003. Preferentially localized dynein and perinuclear dynactin associate with nuclear pore complex proteins to mediate genomic union during mammalian fertilization. Journal of cell science. 116, 4727-38

Pixton, K.L., et al., 2004. Sperm proteome mapping of a patient who experienced failed fertilization at ivf reveals altered expression of at least 20 proteins compared with fertile donors: Case report. Human reproduction. 19, 1438-47

Ramio-Lluch, L., et al., 2014. Oligomycin a-induced inhibition of mitochondrial atp-synthase activity suppresses boar sperm motility and in vitro capacitation achievement without modifying overall sperm energy levels. Reproduction, fertility, and development. 26, 883-97

Rawe, V.Y., et al., 2008. The role of sperm proteasomes during sperm aster formation and early zygote development: Implications for fertilization failure in humans. Human reproduction. 23, 573-80

Reyes, R., et al., 1993. Energy metabolism of spermatozoa during pronucleus formation induced in vitro by heparin-reduced glutathione. I. Glucose uptake. Archives of andrology. 30, 73-7

Rogers, B.J., et al., 1977. Inhibition of hamster sperm acrosome reaction and fertilization by oligomycin, antimycin a, and rotenone. The Journal of experimental zoology. 199, 129-36

Ross, J.M., et al., 2015. Mitochondrial and ubiquitin proteasome system dysfunction in ageing and disease: Two sides of the same coin? International journal of molecular sciences. 16, 19458-76

Segref, A., et al., 2014. Pathogenesis of human mitochondrial diseases is modulated by reduced activity of the ubiquitin/proteasome system. Cell metabolism. 19, 642-52

Sharma, R., et al., 2013. Proteomic analysis of human spermatozoa proteins with oxidative stress. Reproductive biology and endocrinology : RB&E. 11, 48

Shivaji, S., et al., 2009. The role of mitochondrial proteins in sperm capacitation. Journal of reproductive immunology. 83, 14-8

Siva, A.B., et al., 2014. Inhibiting sperm pyruvate dehydrogenase complex and its e3 subunit, dihydrolipoamide dehydrogenase affects fertilization in syrian hamsters. PloS one. 9, e97916

Sousa, A.P., et al., 2011. Not all sperm are equal: Functional mitochondria characterize a subpopulation of human sperm with better fertilization potential. PloS one. 6, e18112

Sutovsky, P., 2011. Sperm proteasome and fertilization. Reproduction. 142, 1-14

Sutovsky, P., et al., 1999. Ubiquitin tag for sperm mitochondria. Nature. 402, 371-2

Swain, J.E., Pool, T.B., 2008. Art failure: Oocyte contributions to unsuccessful fertilization. Human reproduction update. 14, 431-46

Terada, Y., et al., 2010. Essential roles of the sperm centrosome in human fertilization: Developing the therapy for fertilization failure due to sperm centrosomal dysfunction. The Tohoku journal of experimental medicine. 220, 247-58

Torra-Massana, M., et al., 2018. Sperm telomere length in donor samples is not related to icsi outcome. Journal of assisted reproduction and genetics. 35, 649-657
Torra-Massana, M., et al., 2019. Novel phospholipase c zeta 1 mutations associated with fertilization failures after icsi. Human reproduction. 34, 1494-1504
Tosti, E., Menezo, Y., 2016. Gamete activation: Basic knowledge and clinical applications. Human reproduction update. 22, 420-39
Uhlen, M., et al., 2015. Proteomics. Tissue-based map of the human proteome. Science. 347, 1260419
Walsh, A., et al., 2008. Identification of the molecular chaperone, heat shock protein 1 (chaperonin 10), in the reproductive tract and in capacitating spermatozoa in the male mouse. Biology of reproduction. 78, 983-93
Who., 2010. Who laboratory manual for the examination and processing of human semen. 5th edition
Yanagida, K., 2004. Complete fertilization failure in icsi. Human cell. 17, 187-93
Yeste, M., et al., 2016. Oocyte activation deficiency: A role for an oocyte contribution? Human reproduction update. 22, 23-47
Zhang G, et al., 2016 Mitochondrial Biomarkers Reflect Semen Quality: Results from the MARCHS Study in Chongqing, China. PLoS ONE 11(12): e0168823
Zhang, G., et al., 2019. Minocycline impedes mitochondrial-dependent cell death and stabilizes expression of hypoxia inducible factor-1alpha in spinal cord injury. Archives of medical science : AMS. 15, 475-483
Zigo, M., et al., 2018. Modifications of the 26s proteasome during boar sperm capacitation. Cell and tissue research. 372, 591-601

8.

ABREVIATURES I ANGLICISMES

8. ABREVIATURES I ANGLICISMES

A

Àcid – urea PAGE	Electroforesi en gel de poliacrilamida àcid – urea
ACRBP	*Acrosin binding protein*
ADAM	*Disintegrin and metalloproteinase domain-containing protein*
ADAM1B	*Disintegrin and metalloproteinase domain-containing protein 1b*
ADAM2	*Disintegrin and metalloproteinase domain-containing protein 2*
ADAM3	*Disintegrin and metalloproteinase domain-containing protein 3*
ADAM7	*Disintegrin and metalloproteinase domain-containing protein 7*
AKAP3	*A-kinase anchor 3*
AKAP4	*A-kinase anchor 4*

B

Backsplicing	Enllaç covalent entre l'extrem 3' i 5' d'un RNA
***Bottom-up* MS**	Espectrometria de masses basada en pèptids

C

C7ORF61	*Uncharacterized protein C7orf61*
CAMK4	*Calcium/calmodulin dependent protein kinase IV*
cAMP	Monofosfat d'adenosina cíclic
Cargo	Càrrega
CASA	*Computer-Assisted Sperm Analysis*
CatSper	Canals de cations de l'espermatozoide
CATSPER1	*Cation channel sperm-associated protein 1*
CD9	*CD9 antigen*
cDNA	Àcid desoxiribonucleic complementari
circRNA	RNA circular
CNOT1	*CCR4-NOT Transcription Complex Subunit 1*
CNV	*Copy number variation*
CpG	Dinucleòtids citosina-fosfat-guanidina
CRISP1	*Cysteine-rich secretory protein 1*

D

DGCR8	*DGCR8 Microprocessor Complex Subunit*
DICER1	*Endoribonuclease Dicer*
DLAT	*dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex*
DNA	Àcid desoxiribonucleic
DNasa	Desoxiribonucleasa
DNMT	Metiltransferasa
DNMT1	*Methyltransferase 1*
DNMT3A	*DNA (cytosine-5)-methyltransferase 3A*
DNMT3B	*DNA (cytosine-5)-methyltransferase 3B*
DROSHA	*Ribonuclease 3*

E

EEF1A	*Eukaryotic Translation Elongation Factor 1 Alpha 1*
EIF3F	*Eukaryotic translation initiation factor 3 subunit F*
EIF4A2	*Eukaryotic initiation factor 4A-II*
ELP3	*Elongator Acetyltransferase Complex Subunit 3*

F

FIV convencional	Fecundació *in vitro* convencional
FIV-ICSI	Fecundació *in vitro* amb injecció intracitoplasmàtica
FPKM	*Fragments per kilobase of exon model per million reads mapped*
FSH	Hormona estimuladora de fol·licle

G

GnRH	Hormona alliberadora de gonadotropina
GTPasa	Activitat guanosina trifosfatasa

H

H1	*Histone H1*
H1t2	*Testis-specific H1 histone*
H2A	*Histone H2A type 2-A*
H2B	*Histone H2B type 1*
H3	*Histone H3.1*
H4	*Histone H4*
HCl	Àcid clorhídric
HFD	*High-fat diet*
HP2	Forma madura HP2
HP3	Forma madura HP3
HP4	Forma madura HP4
HPI1	*Basic nuclear protein HPI1*
HPI2	*Basic nuclear protein HPI2*
HPS1	*Basic nuclear protein HPS1*
HPS2	*Basic nuclear protein HPS2*
HSPA4L	*Heat shock 70 kDa protein 4L*

I

IE	Element intrònic retingut
IIA	Inseminació intrauterina artificial
IMC	Índex de massa corporal
Imprinting	Impremta gènica
In-gel digestion	Digestió dels pèptids en el gel
In-solution digestion	Digestió dels pèptids directa en solució
INTS1	*Integrator complex subunit 1*
IP_3	Inositol trifosfat
IZUMO1	*Izumo Sperm-Egg Fusion 1*

J

JUNO	*Sperm-egg fusion protein Juno*

K

KIAA0586	*Protein TALPID3*
KO	*Knock-out; genoanul·lat*

L

Label-free	Sense marcatge
LC	Cromatografia líquida
LC-MS/MS	Cromatografia líquida acoblada a espectrometria de masses en tàndem
LEFTY1	*Left-right determination factor 1*
LH	Hormona luteïnitzant
LINE	*Long interspersed nuclear element*
LINE1	*Class 1 long interspersed nuclear element*
***Long* RNA**	RNA de > 200 nt
Loop	Bucle
LTR	*Long terminal repeat*

M

m^6A	Modificació N6-metiladenosina
Microarray	Micro-matriu
miRNA	microRNA
Missing protein	Proteïna predita però no validada experimentalment
mitoRNA	RNA mitocondrial
Mobilitat de tipus A	Espermatozoides amb mobilitat progressiva ràpida
Mobilitat de tipus B	Espermatozoides amb mobilitat progressiva lenta
Mobilitat de tipus C	Espermatozoides amb mobilitat no progressiva
Mobilitat de tipus D	Espermatozoides immòbils
mRNA	RNA missatger
MS	Espectrometria de masses
MTG1	*Mitochondrial Ribosome Associated GTPase 1*
mtRNA	*Small* RNA derivat del mitocondri

N

ncRNA	RNA no codificant
NGS	*Next generation sequencing*
NH	Domini nucleohistona
NP	Domini nucleoprotamina

O

ORF	*Open Reading Frame*

P

P1	Protamina 1
P2	Protamina 2
Paired-end	Seqüenciació pels dos extrems
Paternally-derived **circRNA embrionari**	circRNA embrionari aportat exclusivament per la via germinal paterna
PCR	Reacció en cadena de la polimerasa
PCSK4	*Proprotein convertase subtilisin/kexin type 4*
PGC	*Primordial germ cells*
PIP_2	Fosfatidil-inositol-4,5-bifosfat
piRNA	*P element-induced wimpy testes-interacting (piwi-interacting) RNA*
PIWI	*Protein PIWI*
PKA	*Protein kinase A*
PLC	*Phospholipase C*
PLCζ	*Phospholipase C-zeta*
PLCZ1	*Phospholipase C-zeta-1*
Poli (A)	Poliadenilació
Pool	Conjunt / Barreja de mostres
PPP1CC2	*Serine/threonine-protein phosphatase PP1-gamma catalytic subunit*
Pre-miRNA	Precursor de miRNA
Pre-P2	Precursor de la protamina 2
Primers	Encebadors
Pri-miRNA	Precursor del precursor de miRNA
PRM1	Gen de la protamina 1
PRM2	Gen de la protamina 2
Protein spot	Punt proteic
PSM	*Peptide spectrum match*
PTM	Modificació posttraduccional
PTPRC	*Receptor-type tyrosine-protein phosphatase*

R

RBP	Ribonucleoproteïna
Read	Lectura
Reporter ion	Ió informador
RIN	*RNA integrity number*
RNA	Àcid ribonucleic
RNasa R	Ribonucleasa R
RNA-seq	Seqüenciació massiva d'RNA
ROS	Espècies reactives d'oxigen
RPKM	*Reads per kilo base per million mapped reads*
RPSA	*40S ribosomal protein SA*
rRNA	*Ribosomal RNA*

S

SDS – PAGE	*Sodium dodecyl sulfate polyacrylamide gel electrophoresis*
SINE	*Short interspersed nuclear element*
Single-End	Seqüenciació per un extrem
siRNA	*Small interfering RNA*
***Small* RNA**	RNA de < 200 nt
snoRNA	*Small nucleolar RNA*
SNP	*Single nucleotide polimorphism*
snRNA	*Small nuclear RNA*
SOF1	*WD repeats and SOF1 domain containing*
Somatic cell lysis buffer	Tampó de lisis cel·lular somàtica
SPACA6	*Sperm acrosome membrane-associated protein 6*
SPESP1	*Sperm equatorial segment protein 1*
Splicing	Tall i unió
SRE	*Sperm RNA element*
SRPK1	*SRSF protein kinase 1*
Swim-up	Tècnica de selecció espermàtica

T

Tag	Marca / Etiqueta
TCA	Àcid tricloroacètic
TMEM95	*Transmembrane protein 95*
TMT™	*Tandem mass tag™*
***Top-down* MS**	Espectrometria de masses basada en proteïna intacta
TRA	Tècnica de reproducció assistida
TRBP	*Human immunodeficiency virus transactivating response RNA-binding protein*
tRF	*Transfer RNA fragment*
tRNA	RNA transferència
TRiC	*Chaperonin TCP-1 ring complex*
TRIM66	*Tripartite motif-containing protein 66*
TSSK2	*Testis-specific serine/threonine-protein kinase 2*

U

UniProtKB	*UniProt knowledgebase (https://www.uniprot.org/)*
UTR	*Untranslated region*

V

VRK1	*Serine/threonine-protein kinase VRK1*

W

WHO	*World Health Organization*
WT	*Wild-type*

Y

YRNA	RNA derivat de Y

Z

Zinc finger	*Dit de zinc*
ZP1	*Zona pellucida sperm-binding protein 1*
ZP2	*Zona pellucida sperm-binding protein 2*
ZP3	*Zona pellucida sperm-binding protein 3*
ZP4	*Zona pellucida sperm-binding protein 4*
ZP3R	*Zona pellucida sperm-binding protein 3 receptor*
ZPBP1	*Zona pellucida-binding protein 1*
ZPBP2	*Zona pellucida-binding protein 2*

ALTRES

1D	Monodimensional
2D	Bidimensional
2D-LC-MS/MS	Doble separació per cromatografia líquida seguit d'espectrometria de masses en tàndem
ºC	Graus Celsius / centígrads
µg	Microgram
µm	Micròmetre
bp	Parell de Bases
Da	Dalton
fg	Femtogram
m	Metre
m / s	Metre / segon
m / z	Massa / càrrega
ml	Mil·límetre
ng	Nanogram
nt	Nucleòtid
kb	Quilobase