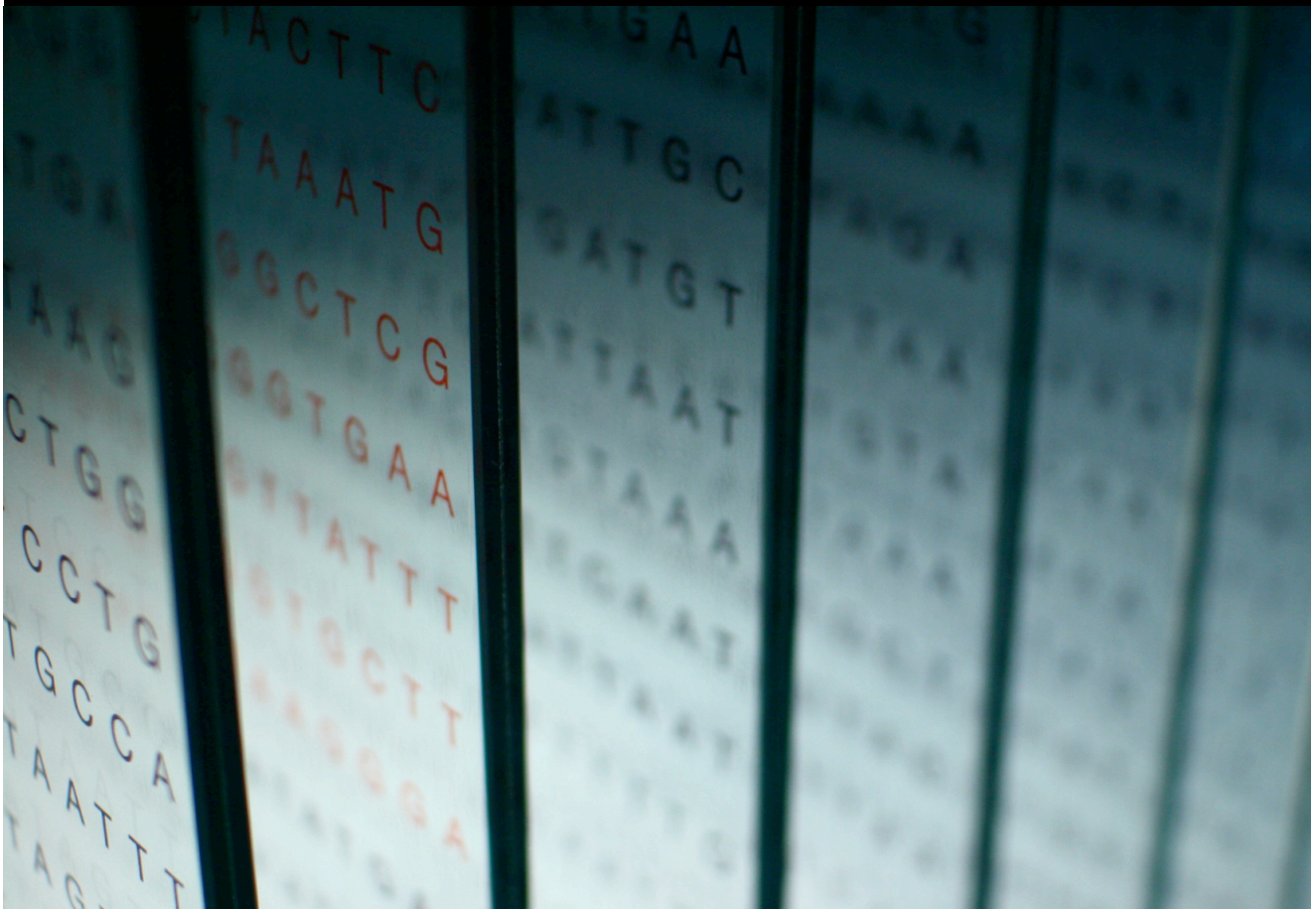


Tesi Doctoral

Alteracions epigenètiques en càncer colorectal

Canvis globals i identificació
de noves dianes



Jairo Rodríguez Lumbarres
Novembre 2007

Alteracions epigenètiques en càncer colorectal: canvis globals i identificació de noves dianes

Memòria presentada per
Jairo Rodríguez Lumbarres

per optar al grau de
Doctor en Biologia

Tesi realitzada sota la direcció del Dr. Miquel Àngel Peinado
a l'IDIBELL-Institut de Recerca Oncològica.

Tesi adscrita al Departament de Genètica de la Facultat de Biologia
Universitat de Barcelona.
Programa de Genètica, bienni 2003-2005.

Director

Tutor

Miquel À. Peinado

Ricard Albalat

Jairo Rodríguez

Barcelona, Novembre de 2007

A la Laura

Als meus pares

Al meu germà

Agraïments

Tot arriba. Fins i tot allò que un dia veiem molt llunyà acaba per atrapar-nos. Ja deu de fer més de 5 anys que corro pels laboratoris amb tots vosaltres. Alguns m'heu acompanyat en aquest viatge de més a prop, i segurament heu pogut veure la meva evolució durant aquest temps, al mateix temps que jo he pogut veure els canvis que s'han produït en vosaltres. Ha arribat el meu torn. Ara em toca a mi emprendre una nova direcció i no puc sinó sentir nostàlgia per totes aquelles coses que hem viscut junts. He après infinitament amb vosaltres i m'heu ajudat a créixer com a persona i com a científic, si es que em puc qualificar d'això últim. De fet aquests agraïments són una mica en va. No us ho puc agrair amb paraules, sinó amb l'ofertament de la meva amistat sincera. És igual on vagi, porto una maleta gran, i aquesta etapa i tots vosaltres m'hi acompanyareu sense remei.

He de començar per tu, MAP. No oblidaré mai el dia que vaig rebre la trucada en la que em deies que podia unir-me al teu grup. Passava junt a l'Alcampo de Sant Quirze amb els "Ferrocatas" i quasi em poso a saltar i ballar al mig del vagó. Aleshores encara tenia vergonya. La teva personalitat excepcional m'ha fet entendre que encara hi ha persones que no es mouen per l'ambició. Gràcies per tenir la porta del teu despatx sempre oberta i sobretot per deixar-me volar quan ho necessitava. De fet, podem dir que tots dos hem volat, encara que algun que altre aterratge fos una mica més dolorós, com la vida mateixa no? Espero que trobis allò que busques en aquesta nova etapa que has decidit emprendre.

Eli, la meva primera mestra en el laboratori. Amb tu vaig començar a sentir la màgia de veure com les bandes apareixen en els gels de plata i encara avui recordo la sensació que se sent quan després de tanta feina apareix davant dels teus nassos un bon patró de bandes! La teva tranquil·litat i sinceritat són quasi d'un altre món i sempre has tingut bons consells per a mi.

Jordi, per (a), per (b) o per (c), tu i jo no hem coincidit mai gaire, ni en el temps ni en l'espai, però tot i això crec que he pogut aprofitar al màxim els moments viscuts plegats, ja sigui discutint en el lab o caminant per algun paradís perdut. No et puc agrair amb paraules l'haver compartit amb mi la teva especial manera de veure el món. Jo no sóc tant pragmàtic, ni molt menys, però em vas fer veure en el moment que més ho necessitava que cadascú ha de fer el seu camí. I tu de camins n'has recorregut uns quants, inclús aquells en els que el deliri momentani provocat per la indescriptible bellesa del paisatge et fa al·lucinar i tot d'una veus noies lleugeres de roba. Perquè va ser una al·lucinació, no?

Mar, de veritat ets incansable? A tu no et puc agrair res en concret, seria una llista massa llarga. Quanta feina feta i quantes converses i discussions; quantes festes i quantes enrabiades i quants moments de felicitat quan les coses surten com un vol. Curiosament la meva primera escapada a Seattle ens va apropar i des d'aleshores aquesta tesi no existiria sense tu. Ets un llibre obert de com s'haurien de fer les coses en un món perfecte, encara que potser en un món perfecte l'home no hi seria. Sigui com sigui, no canviïs la teva filosofia de vida, encara que de vegades adaptar-nos a les noves situacions faci la vida una mica més fàcil. I no et creguis mai que tu no pots arribar allà on et proposis, només és qüestió de provar-ho!

¿Y dónde está Cris? ¡En cultivos! Bueno, en cultivos o en el citovision. Empezamos casi a la vez y ya te tenemos en Dinamarca. Gracias por tus consejos y conversaciones de escalera y sobre todo por amenizarnos las tardes con tus monólogos...tranquila, yo también hablo conmigo mismo, aunque más bajito.

I la Gemma, Gemícola! Sempre seré el teu Janairo! Ai, quantes coses hem viscut plegats en el Map Lab. Sembla que va ser ahir quan la Laia va nàixer i jo vaig tenir la meva primera feina! Gràcies per tots els riures que ens hem fet plegats i per la teva ajuda i sinceritat sempre que l'he necessitat!!

Y Regi. A veces hemos estado más cerca y otras veces más lejos, pero debo agradecerle la energía y las ganas de seguir adelante que siempre contagias a todo el mundo. A veces debería seguir un poco más tu positivismo, seguro que las cosas serían un poco más fáciles.

A la biòloga de camp, l'Elena. La teva força i determinació per la feina són admirables. Gràcies per deixar-me compartir amb tu nits de Baldriga i cafè sobre una pedra del Mediterrani. Tot i que el temps i la distància ens han allunyat guardo amb molt bon record els dies en els que la meva companya de poiata jugava als vampirs amb tota mena de bestioles.

A la Maria, l'Eivissenca esbojarrada. Les circumstàncies no van voler que coincidíssim molt de temps en el laboratori, però les meves escapades ciclistes a Eivissa em van permetre de conèixer-te una mica millor. La teva mirada crítica i directa no sempre és compartida, però proporciona moments de discussió que val la pena no perdre's.

A la Mire, la muntanyenca del Map Lab. Saps que fas molta enveja (de la de bona fe) amb la teva furgo? Gràcies per totes aquelles xerrades d'escalada, bicis, trekkings i demés que deixen descansar una mica el cap entre les quatre parets del lab. I sobretot molta força i ànims en aquesta nova etapa que has començat, ara sí, una mica més a prop de casa teva.

A les noves incorporacions del Map Lab, la Marta i la Inês i la Berta. És genial veure com per fi arriba algú més jove que jo al lab! Per fi ja sóc gran! M'ha agradat molt compartir aquest temps, encara que curt, amb vosaltres. Us desitjo molta sort i sobretot, passeu-ho bé en el camí, que no se sap mai que hi ha al final.

A tú, Antònia. Moltes gràcies pels teus consells del dia dia, sempre has tingut una resposta per mi, encara que de vegades m'hagi endut alguna que altra bronca! I la resta de nenes del Fabra lab, les que encara hi són i les que ja no, Anna, Eva, Marinela, Olga, Cristina i Àngels. El nostre desplaçament a l'ICO no va fer que deixés de fer-vos visites per fer intercanvi de reactius. Gràcies per la vostra ajuda i amistat.

Anna i Lluís. Gràcies per les discussions científiques i idees que heu aportat en aquesta etapa final de la meva tesi. Els meus primers ChIPs a Barcelona van ser una mica menys difícils gràcies a vosaltres!

I a la resta dels membres del COM, als que hi heu estat des de bon principi, als que us heu anat incorporant durant el camí i als que ja no hi sou. Olga, Laura, Ruth, Sergio, Quim, Bernat, Sergi, M^a Àngels, Violeta, Eder, Helena, Rebeca, Àlex, Vane, Vero, Cris, Cristina, Ana, Conrad, Miguel, Jaume... i segur que em descuido algú! (perdó!!). Heu fet del COM el departament més alegre i divertit de tot l'IRO! Gràcies per tots els "bons dies" que heu tingut per a mi i els bons moments que hem passat plegats, incloses les comilonas del COM!

I com no, als nostres veïns i acollidors a l'ICO. Josep Maria, Álvaro, David i Alberto. És increïble la de conyes que se senten des del nostre (ja ex-) laboratori! Gràcies per la vostra amistat i totes les festes que hem viscut plegats. Josep Maria, gràcies per tot el suport i amistat durant aquests anys i algun que altre concert elèctric. Gràcies a tu també David per compartir amb mi la teva passió per la fotografia, algun dia faré fotos com tu. I a tu Álvaro, llàstima que la nostra vida tranquil·la de becaris s'acabés, i amb ella les tardes d'escalada a Garraf i alguna que altra sortida per Collserola. Cada cop que surto passo per la trialera que em vas descobrir! Alberto, a ti tengo que agradecer que me enseñaras a jugar con los gusanitos. Fue divertido mientras duró. Gracias por tus consejos, personales y científicos. Antònia i Mireia, amb vosaltres vàrem compartir migdiades als seminaris de recerca i ara em toca a mi el torn de llegir. Gràcies per la vostra amistat i converses de passadís.

A la resta de membres de l'ICO, Gabi, Ramón, Agnés, Mònica, Olga, Felip, Paco, Marta, Neus, Vane, Raquel, Ernest, Mireia, Dani, Laura C., Jéssica, Glòria, Laura P., Maria, Marc...Gràcies a tots aquells que en un moment o altre m'heu ajudat i heu compartit amb mi la vostra manera de ser. I a la gent de l'ICO amb qui hem compartit lab, Nadia, Bego, Fati, Xavi i Núria. Ha estat un plaer compartir les poiates amb vosaltres. Y...iay Isa! Gracias por tu simpatia y desparpajo...iia no te podré pisar más, el suelo digo!!

Gràcies també a tu Eduard. Sempre amb una pila de preguntes a punt al final d'un seminari. Irradies energia al teu voltant i encara que t'he conegut tard, m'has donat molt bons consells per al futur. I a tu Enric, gràcies per compartir amb mi moments de rock. No arribarem mai gaire lluny, però hem passat bones estones de guitarreo. Marga, moltes gràcies pels ànims que sempre m'has donat i les converses al quartet del cito!

Gràcies també a tu Ricard, per haver-me donat suport sempre que t'he necessitat, sobretot en la recta final de la meva tesi.

How could I forget my scientific family overseas? Sincerely thanks Mark for having me in your lab and let me play around for that long. My stay in Seattle was an accelerated course of science and life. And thank you Gus for your unconditional support and ideas. This thesis wouldn't be as it is without you. Héctor, ¡pinche Mexicano! Qué suerte fue conocerte. No podré olvidar aquel tiempo en Seattle con el mexicano más loco. Muchas fiestas y mucho trabajo y pelis de superhéroes me ayudaron a pasar un tiempo un poco lluvioso en la Emerald City. Gràcias por la amistad que me ofreciste y que todavía sigue. Jessica, Agnes, Rachel, Stephy, Nanna, David, Tim, Tomo and Tobias: thank you so much for all your help and friendship, you really made me feel at home in the Groudine Lab! And thank you Steve for letting me share with you some of the best philosophical moments, always with some good beer.

Carles, Jéssica i Adrià. Gràcies per tota l'ajuda que m'heu ofert en aquesta etapa final de la meva tesi. Ha estat un plaer conèixer-vos!

I a la colla del SingStar! Sílvia, Xavi, Uri, Jaume, Edu, Vicenç, Joan i Gemma. Gràcies per haver compartit amb mi nits de galls i divos. Siguem realistes nois, encara que ens hi esforcem molt només podem ser superestrelles en mode fàcil, tot i que unes herbes facin miracles! Les nits de cap de setmana amb vosaltres són el millor antiestrés!

Als meus amics, Jose i Ubald. Jose, hemos vivido muchas cosas juntos desde que nos conocimos. Hemos caído, andado, bailado i reído mucho desde entonces. Sobre todo nos acordamos de los porrazos a dos ruedas ¿no? Aunque ahora ya no caemos juntos, todavía reímos y bailamos y ihasta nos hemos casado! De ti sí que me lo esperaba, pero ¿yo? I tu Ubald! Qui m'anava a dir que compartírem moments tant importants de la nostra vida plegats? No et preocupis, no hi haurà més llistes. Us agraeixo als dos més enllà de les paraules la vostra amistat, tot i les meves absències físiques i mentals més o menys duradores. Gràcies per ser-hi.

Al meu germà, Jordi. M'alegro de que en el nostre cas la distància física ens hagi apropat més que mai. Gràcies per deixar que et conegui una mica més cada dia i pel teu suport sense condicions. Espero poder fer molts més vídeos inútils de dos germans en bicicleta i moltes més sortides fotogràfiques d'aficionat.

Als meus pares, Esther i Paco. A vosaltres us dono les GRÀCIES per haver-me permès ser la persona que sóc. Gràcies per haver-me comprat el microscopi, els jocs de química i els atlas de biologia quan encara no sabia que era el DNA. Gràcies per haver-me ensenyat a valorar les coses i a lluitar pel que es desitja i sobretot, gràcies per haver-me donat l'oportunitat de començar a viure la vida lluny de casa i estudiar allò que tant desitjava. Gràcies per creure que allò que deia un nen de deu anys, que volia ser científic, es podria convertir en realitat.

I finalment, a tu. Com agrair-te la llum que dones a la meva vida? Tu ets la meva companya de viatge, també d'aquest viatge que ara acabo, la persona que em dona la mà quan me'n vaig a terra i la persona amb la que descobreixo el món cada dia. Ets part de mi i no m'espanto al pensar en tots els reptes que tenim al davant. Sempre diem que Hawai'i va ser el principi d'una bona època, però jo crec que la bona època va començar el dia que vàrem decidir caminar plegats. Ja saps el que diu el magnet de la nevera, així que no pensem en el demà i disfrutem al màxim avui.

I ja no m'estenc més. Només dir que pels que em coneixeu, ni que sigui una mica, sabeu de la meva passió per la bicicleta. Un dia, farà ja 15 anys, estava fullejant una "Solo Bici" en la que hi havia un especial sobre com afrontar les pujades. L'especialista en qüestió, de nom Ned Overend, feia una reflexió que m'ha acompanyat des d'aleshores.

Les pujades sempre fan mal; a mesura que ens fem més bons, simplement anem més ràpid.

Índex de continguts

Abreviacions

Relació de figures

Relació de taules

Pròleg

Introducció

1. Què és el càncer?	1
2. Inestabilitat genètica	6
3. Epigenètica	9
4. Alteracions epigenètiques en càncer	34
5. Tècniques d'anàlisi de la metilació	44

Objectius	51
-----------	----

Resultats

Capítol 1. <i>Chromosomal instability correlates with genome-wide DNA Demethylation in human primary colorectal cancers</i>	55
Capítol 2. <i>Genome-wide tracking of unmethylated DNA alu repeats in normal and cancer cells</i>	67
Capítol 3. <i>Bivalent domains enforce transcriptional memory of DNA methylated genes in colorectal cancer.</i>	117

Discussió 161

Conclusions 185

Bibliografia 189

Apèndix 213

Abreviacions utilitzades

AIMS:	amplification of intermethylated sites
AP-PCR:	arbitrarily primed polymerase chain reaction
AUMA:	amplification of unmethylated Alu
BET:	bromodomini i ET domini
CCR:	càncer colorectal
ChIP:	chromatin immunoprecipitation
CIMP:	CpG island methylator phenotype
CIN:	chromosomal instability
COBRA:	combined bisulfite restriction analysis
CR:	caloric restriction
CSC:	cancer stem cell
DNMT:	DNA metiltransferasa
DNA:	desoxyribonucleic acid
ESC:	embryonic stem cells
EST:	expressed sequence tag
GFP:	green fluorescent protein
HAT:	histone acetyl transferase
HDAC:	histone deacetylase
HKMT:	Histone K (lysine) methyl transferase
HNPCC:	hereditary nonpolyposis colorectal cancer
HPLC:	high performance liquid chromatography
ICF:	immunodeficiency, centromeric instability and facial anomalies
ICR:	imprinting control region
LINE:	long interspersed nucleotide element
LOH:	loss of heterozygosity
LOI:	loss of imprinting
Mb:	megabase; milió de parells de bases

MBD:	methyl binding domain
meDIP:	methylated DNA immunoprecipitation
MIN:	microsatellite instability
mRNA:	messenger RNA
MS-AFLP:	methylation-sensitive amplified fragment length polymorphism
MSI:	microsatellite instability
NAD:	nicotinamida adenina dinucleòtid
Pb:	parell de bases
Pc:	polycomb
PcG:	polycomb group
PCR:	polymerase chain reaction
PEV:	position-effect variegation
PIC:	preinitiation complex
PRC1/2:	polycomb repressor complex 1 and 2
PRE:	polycomb response element
PTM:	posttranslational modification
QUAMA:	quantification of unmethylated Alu
RdDM:	RNA directed DNA methylation
RLGS:	restriction landmark genome scanning
RNA:	ribonucleic acid
RNAi:	RNA d'interferència
SAM:	s-adenosil-metionina
SINE:	short interspersed nucleotide element
siRNA:	small interfering RNA
TAF:	transcription associated factor
TRD:	transcriptional repressor domain
Trx:	trithorax
TSS:	transcription start site
5mC:	5 metil citosina

Relació de figures

Figura 1. Anatomia del tram final de l'aparell digestiu humà

Figura 2. Hipòtesi revisada de les dues alteracions de Knudson

Figura 3. Model de progressió del càncer colorectal

Figura 4. Les 6 capacitats adquirides de les cèl·lules tumorals

Figura 5. La variabilitat genètica és el motor de la tumorigènesi

Figura 6. Evolució en el nombre de publicacions sobre epigenètica en els darrers anys

Figura 7. Del DNA al cromosoma metafàsic

Figura 8. Estructura del nucleosoma

Figura 9. Estructura de la citosina i la 5 metil citosina

Figura 10. Les illes CpG

Figura 11. Les modificacions de les histones

Figura 12. Seqüència d'esdeveniments genètics i epigenètics en l'activació del gen IFN- β

Figura 13. El codi de metilació de les histones

Figura 14. Diàleg entre els diferents components de la regulació epigenètica

Figura 15. Principals alteracions epigenètiques en càncer

Figura 16. Origen de les seqüències Alu

Figura 17. Estructura genètica de la seqüència Alu

Figura 18. Tractament de DNA amb bisulfit sòdic

Figura 19. Metodologies d'anàlisi de la metilació del DNA

Figura 20. Hipometilació i alteracions cromosòmiques.

Figura 21. El compartiment de cèl·lules mare epitelials en la cripta del còlon.

Relació de taules

Taula 1. Contingut de 5mC en diversos organismes model

Taula 2. Tipus de modificacions covalents de les histones

Taula 3. Classificació de HATs conegudes i putatives

Taula 4. Classificació de HDACs conegudes en llevat i mamífers

Taula 5. Proteïnes que contenen bromodominis i la seva funció

Taula 6. Principals components dels complexos polycomb PRC1 i PRC2

Taula 7. Classificació de les diferents activitats metiltransferasa i desmetilasa conegudes

Taula 8. Gens hipermetilats en càncer

Pròleg

De vegades les comparacions ens permeten entendre millor fenòmens complexos. Pensem doncs en el DNA que hi ha en el nucli com a una gran biblioteca, en la què hi ha diferents prestatges que llisquen sobre rails. En aquesta immensa biblioteca hi ha determinats prestatges que es troben àmpliament separats dels altres, permetent que els lectors que busquen un determinat llibre s'hi puguin passejar i fer un cop d'ull als diferents volums. Hi ha tant d'espai que varis lectors poden estar-s'hi alhora. D'altres prestatges però, es mantenen molt junts entre ells i els lectors no hi poden accedir. Ara imaginem que necessitem un volum que no trobem, i que segurament es troba entre els prestatges inaccessibles. Només farà falta que el bibliotecari separi els prestatges i així podrem accedir a molta més informació que ens era oculta. A la inversa, potser una de les zones de la biblioteca ja no és necessària i pot tancar-se al públic. Si la nostra biblioteca és petita de seguida sabrem mantenir-ne l'ordre, però si disposem de 30.000 volums no és difícil imaginar que ens farà falta un codi que ens recordi quins prestatges han de mantenir-se ocults i quins han de mantenir la seva informació accessible.

Aquesta tesi versa sobre això, sobre els mecanismes que regulen l'accessibilitat a la informació que conté el DNA dins el nucli, a través de la regulació epigenètica. Així, si la informació genètica la trobem dins dels llibres, la informació epigenètica la trobarem en els llocs dels llibres i sobre els prestatges, de manera que pugui ser fàcilment visible i interpretable. Aquesta informació epigenètica configura un veritable codi que disposa quines regions de la informació genètica s'han de mantenir accessibles i quines no. Com veurem més endavant, la cèl·lula disposa del seu conjunt de bibliotecaris particular, els quals són capaços de moure els prestatges (histones) que empaqueten els llibres (DNA). Donat que totes les cèl·lules del nostre cos, encara que agrupades en tipus cel·lulars diferenciats, contenen la mateixa informació o codi genètic (a excepció és clar dels limfòcits B), és imprescindible per al manteniment de la diferenciació que cada tipus cel·lular tingui accés restringit a la informació.

Quan els mecanismes que mantenen aquesta regulació fallen, el conjunt d'informació que una cèl·lula té a la seva disposició canvia, tot perdent informació que li era necessària i guanyant-ne d'altra que li confereix noves capacitats. Les cèl·lules perden la seva identitat i poden deixar de comportar-se d'acord amb les necessitats del teixit en el què es troben, tot conduint a fenòmens com el càncer.

Aquesta tesi es presenta com a un compendi de publicacions i consta d'una introducció i discussió generals i tres treballs, el primer dels quals està ja publicat, un segon està en vies d'acceptació i finalment un tercer treball està llest per ser sotmès per a la seva publicació.

En primer lloc, una introducció general esmenta tots aquells conceptes que s'aniran tocant al llarg de la tesi, seguint més aviat un ordre lògic que connecta els diferents temes més que no pas un ordre estrictament jeràrquic. Al mateix temps, no hem aprofundit per igual en tots els conceptes, sinó que fem especial èmfasi sobre aquells que són més rellevants per aquesta tesi.

A través dels diferents treballs que presentem, ens hem proposat aprofundir en el coneixement dels mecanismes epigenètics que tant freqüentment es veuen involucrats en els processos tumorals, anant des dels estudis més globals als estudis més específics. Així, ens hem centrat

bàsicament en les dues alteracions epigenètiques més rellevants per al procés tumoral: la hipometilació global i la seva relació amb la inestabilitat genètica i la hipermetilació d'illes CpG i la seva relació amb el silenciament transcripcional. Hem de dir que si bé la metilació del DNA ha tingut un pes molt important en l'inici d'aquesta tesi, aquest pes s'ha anat reduint a mesura que hem entrat en l'estudi de regions específiques, tant hipometilades com hipermetilades, en favor de l'estudi de les modificacions de les histones. Per tant, hem pogut arribar a un anàlisi de les alteracions epigenètiques en càncer a nivell més global, més enllà de la metilació del DNA.

Tot i que la hipometilació global i regional del DNA va ser la primera alteració epigenètica detectada en tumors, les dificultats en el seu estudi, així com la manca de metodologies específiques que en permetin un anàlisi detallat, han fet que desconeguem en gran mesura els efectes que aquest fenomen té per a la cèl·lula tumoral. No ha estat fins els darrers anys que estudis en diferents organismes model han començat a aportar evidències d'un paper de la hipometilació del DNA en la inducció d'inestabilitat genètica. Nosaltres ens vàrem proposar realitzar estudis globals que permetessin establir possibles associacions entre la hipometilació i la inestabilitat genètica en tumors primaris humans, la qual cosa no s'havia fet mai.

Així, en el primer treball, la caracterització dels patrons de metilació globals d'una sèrie de tumors colorectals ens va permetre observar una correlació positiva entre el grau de desmetilació dels tumors primaris i el nombre d'alteracions cromosòmiques que aquests acumulen i vàrem poder aprofundir una mica més en l'estudi de les conseqüències que la hipometilació té per al genoma de la cèl·lula tumoral.

Per a continuar aprofundint en l'estudi de la hipometilació en tumors, i donada la manca de metodologies específiques per a realitzar aquesta mena d'estudis, ens vàrem proposar desenvolupar una nova metodologia d'anàlisi de la metilació, centrada en l'anàlisi dels elements repetitius dispersos Alu, els quals contenen una fracció important de la metilació del DNA en humans. Així es com va sorgir la nova metodologia AUMA, recollida en el segon treball, la qual és presenta més com una innovació conceptual que no pas metodològica. L'aplicació d'AUMA a una sèrie de tumors colorectals ens ha permès identificar elements repetitius Alu desmetilats al llarg de tot el genoma i també ens ha permès identificar elements Alu amb alteracions de la metilació en tumors. Un anàlisi bioinformàtic paral·lel ens ha permès complementar els resultats obtinguts experimentalment, fet que ens ha permès aprofundir en l'estudi de la hipometilació en elements Alu tant en teixit normal com tumoral. En aquest treball hem desenvolupat també una nova metodologia de quantificació d'elements Alu desmetilats, la qual hem anomenat QUMA i de la qual no hi ha cap precedent en la literatura.

De l'estudi d'una de les alteracions de la metilació detectades en patrons AUMA n'ha sorgit el tercer treball d'aquesta tesi. En aquest cas hem fixat la nostra atenció en un petit element MIR desmetilat en teixit normal, el qual s'hipermetila en tumors en una elevada freqüència. L'estudi de l'estat de metilació de les illes CpG dels gens adjacents ens ha conduït a la identificació de 8 noves illes CpG hipermetilades en tumors i línies cel·lulars de càncer colorectal. L'anàlisi de les modificacions d'histones associades al silenciament d'aquests gens ens ha permès identificar patrons de cromatina molt semblants als descoberts recentment en cèl·lules mare embrionàries, connectant així la biologia de les cèl·lules mare amb el desenvolupament del càncer.

En darrer lloc, una discussió global debat, més enllà de les discussions pròpies de cada treball, la rellevància dels resultats obtinguts en el context actual. Aquesta discussió segueix en certa manera l'ordre cronològic en el qual hem obtingut els diferents resultats, amb un especial èmfasi sobre els resultats obtinguts en el darrer treball.

INTRODUCCIÓ

1. Què és el càncer?

La paraula càncer té el seu origen en la paraula llatina *cancris*, la qual vol dir cranc. Curiosament, la paraula en grec antic per designar cranc és *καρκίνος* (karkínos), la qual està relacionada amb *καρκίνωμα* (karkínoma). Aquesta última va ser la paraula que Hipòcrates, ja en el segle V a.C., usava per referir-se a qualsevol lesió cancerosa. La similitud que els antics grecs i llatins varen veure entre els les projeccions dels tumors sobre els teixits sans i la morfologia dels crancs, va donar nom a aquesta malaltia (Varis autors 2002).

Avui en dia, càncer és un terme genèric que agrupa tot un conjunt de malalties que tenen com a denominador comú un creixement cel·lular descontrolat més enllà dels seus límits normals, el qual dóna lloc a la formació d'una neoplàsia (que literalment significa "nou creixement") o tumor maligne, que pot afectar virtualment a qualsevol part de l'organisme.

El càncer colorectal

En el cas del càncer de còlon i recte, també anomenat càncer colorectal (abreviat CCR), ens referim a afectacions que s'originen en aquestes dues porcions de la part final del tub digestiu (Figura 1). Tot i que es tracta de dues malalties que afecten dues regions anatòmiques diferents, donada la similitud d'ambdues patologies, en la pràctica són tractades com a una de sola.

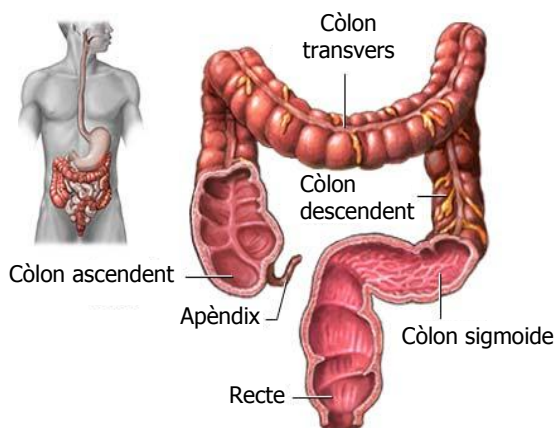


Figura 1. Anatomia del tram final de l'aparell digestiu humà. Aquest el forma principalment el còlon, el qual està separat del intestí prim per la vàlvula íleo-cecal i es pot subdividir en quatre porcions anatòmicament diferenciades anomenades còlon ascendent, transvers, descendent i sigmoide. El recte és la última porció del tub, entre el còlon sigmoide i l'anus.

Epidemiologia del càncer colorectal

Segons dades de la Organització Mundial de la Salut (www.who.org) referents a l'any 2005, el càncer és la segona causa de mortalitat en els països desenvolupats, només per darrera de les malalties cardiovasculars i els accidents. Les dades referents a Espanya mostren que el CCR és la segona causa de mortalitat per darrera del càncer de pulmó en homes i darrera del càncer de mama en dones. A Catalunya, les xifres no són molt diferents de les de la resta d'Espanya i la resta de països desenvolupats. Tal i com es recull en el Pla Director d'Oncologia a Catalunya, entre els anys 2001 i 2004 la taxa d'incidència de CCR en homes i dones va ser molt similar, el

14.6% i el 15.2% respectivament. Les taxes de mortalitat varen seguir una tendència molt similar, essent del 12% per als homes i del 15% per a les dones. La tendència actual, tant a nivell d'Espanya com en la majoria de països desenvolupats, és d'increment del nombre anual de casos de CCR, fet que s'ha atribuït en part a un envelliment progressiu de la població, hàbits de vida més sedentaris i hàbits alimentaris inadequats.

Diagnòstic i tractament del càncer colorectal

El diagnòstic temprà del CCR, quan aquest està en fases molt inicials i localitzades, és de vital importància, donat que hi ha poques estratègies curatives que realment siguin efectives en els casos més avançats de la malaltia. Per aquesta raó s'han desenvolupat estratègies de cribatge en aquells grups de població que tinguin un major risc de desenvolupar la malaltia. Aquests grups de risc inclouen a les persones majors de 50 anys; aquelles que tenen antecedents personals o familiars de CCR o poliposi i aquelles persones amb malaltia inflamatòria intestinal. En aquests grups de risc es poden aplicar diferents mètodes, invasius i no invasius, que permeten detectar directa o indirectament la presència de pòlips i/o càncer.

Els mètodes diagnòstics emprats de forma rutinària en la detecció del CCR són el test no invasiu FOBT (acrònim de l'anglès *faecal occult blood testing*) el qual és un mètode no invasiu enfocat a la detecció de sang en femta, i les endoscòpies (mètode invasiu) de còlon (colonoscòpia) i del còlon sigmoide (sigmoïdoscòpia). Si bé el test FOBT és només una indicació indirecta de la possible presència de pòlips o càncer, les endoscòpies del còlon permeten detectar directament aquesta presència, podent recollir mostres dels teixits afectats per a poder dur a terme un diagnòstic més acurat en el laboratori.

Pel que fa al tractament, la ressecció quirúrgica del tumor continua essent un dels mètodes més eficaços per eliminar el màxim de massa tumoral de l'organisme, fet que permet assolir taxes de supervivència lliure de malaltia a 5 anys del 30-40%. D'altra banda, donada la impossibilitat de saber amb certesa que no queden restes de cèl·lules canceroses després de la cirurgia, tractaments posteriors amb agents químics o radiològics permeten eliminar, de forma més o menys selectiva, possibles restes de cèl·lules tumorals. L'èxit dels mètodes quirúrgics depèn en gran mesura del fet que la massa tumoral no hagi envaït els teixits que l'envolten, es a dir, que es mantingui com a una massa localitzada. D'aquesta manera ens assegurem una extracció de tot el tumor causant el mínim dany possible als teixits que l'envolten.

Si tenim en compte però el gran nombre de pacients de CCR que recauen després d'haver estat tractats per primer cop, es a dir, que pateixen recidiva, queda clar que els mètodes actuals de tractament no són tot l'efectius que caldria esperar. Per tant queden per millorar les tècniques que permetin un diagnòstic més precoç de la malaltia quan aquesta és més fàcilment curable i en segon lloc, millorar els mètodes terapèutics.

El càncer colorectal com a malaltia genètica en múltiples fases

El fet que el procés neoplàsic resulta de l'acumulació progressiva d'alteracions genètiques queda avui en dia fora de tot dubte, sustentat per estudis moleculars, epidemiològics i citogenètics duts a terme al llarg de quasi un segle. En aquest context, els tumors colorectals representen un sistema excel·lent per a la recerca i estudi d'aquestes alteracions donat l'extens coneixement que tenim sobre la seva història natural en múltiples fases, el seu desenvolupament lent i la facilitat que tenim en l'obtenció de mostres en diferents estadis. Fent un resum molt simplista, la recerca realitzada fins ara demostra que el CCR resulta, en part, de l'acumulació progressiva de mutacions en oncogens i gens supressors tumorals (Revisat en (Fearon and Jones 1992)), la qual cosa permet a una cèl·lula de l'epiteli del còlon o el recte trencar l'equilibri que manté amb les cèl·lules veïnes. Això ho aconsegueix per mitjà del guany de funcions que no tenia o tenia limitades i la pèrdua de funcions que li eren pròpies i que en general, van destinades a mantenir la cèl·lula quiescent i diferenciada. Els guanys de funció solen anar associats a la mutació d'oncogens, essent necessària en aquest cas l'alteració d'una única còpia del gen (mutació dominant). Per contra, les pèrdues de funcions cel·lulars solen anar associades a mutacions en gens supressors essent en aquest cas necessàries mutacions en les dues còpies del gen per a la supressió total del mateix (mutació recessiva).

El fet que són necessàries dues mutacions per tal d'inactivar un gen amb funció supressora de tumors ja va ser hipotetitzat per Knudson vint anys enrere (Revisat en (Knudson 2001)) i té en l'actualitat una àmplia acceptació. En el model original però, només es reconeixien dos mecanismes capaços d'inactivar l'activitat d'un gen: la ja esmentada mutació i la pèrdua d'heterozigositat o LOH (acrònim de l'anglès *Loss of Heterozygosity*). La LOH consisteix en la pèrdua física de seqüència de DNA, que pot anar des dels pocs milers de bases i involucrar pocs gens fins a regions cromosòmiques que engloben milions de parells de bases. La recerca realitzada en els darrers anys ha permès descobrir un tercer mecanisme molt important de inactivació de la funció gènica el qual discutirem de forma extensa en l'apartat d'epigenètica, que és el silenciament transcripcional associat a la hipermetilació de DNA en regions promotores. Donat que tots tres mecanismes tenen la mateixa conseqüència, la inactivació d'un al·lel d'un gen, seria lògic pensar que qualsevol combinació dels tres mecanismes es dona en la realitat dels tumors. L'evidència experimental però, ha demostrat que per a cada gen en concret i en cada tipus de tumor, determinades combinacions prevalen sobre les altres (Figura 2) (Revisat en (Jones and Baylin 2002)).

Molt freqüentment els avenços en el coneixement van lligats als avenços en la tecnologia. En aquest sentit, el gran desenvolupament de les tècniques d'anàlisi molecular durant la dècada dels 80 i principis dels 90 varen permetre un estudi més profund del repertori d'alteracions moleculars en tumors, començant a traçar el camí mutacional que guia a una cèl·lula normal cap a la tumorigènesi. De forma molt important, els treballs duts a terme per Vogelstein i col·laboradors a finals dels 80 varen ser pioners en l'elaboració del primer model de progressió de CCR basat en dades moleculars (Vogelstein et al. 1988).

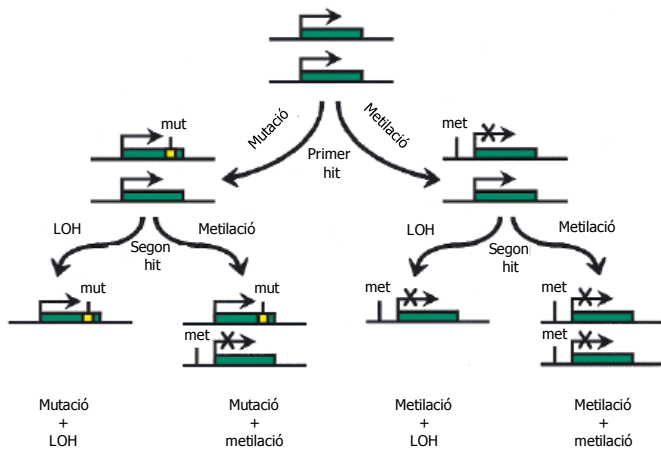


Figura 2. Hipòtesi revisada de les dues alteracions de Knudson. En l'esquema es mostren les diferents combinacions descrites dels tres mecanismes reconeguts d'inactivació de l'activitat gènica en càncer. Com es pot comprovar, rarament es presenta una pèrdua d'heterozigositat (LOH) com a primer esdeveniment o "hit". És interessant fer notar que la metil·lació és l'únic mecanisme que afecta més freqüentment als dos al·lells d'un mateix gen. Modificat de (Jones and Laird 1999).

El model original, el qual podem veure en forma de diagrama en la figura 3, ha estat objecte de revisions posteriors (Fearon and Jones 1992; Fearon and Vogelstein 1990; Kinzler and Vogelstein 1996; Toribara and Sleisenger 1995). La importància d'aquest model rau en el fet que per primer cop es va mostrar que existeix una certa correlació entre l'acumulació progressiva d'alteracions moleculars i les diferents fases de progressió tumoral a partir de teixit normal passant per l'estadi d'adenoma, carcinoma i finalment metàstasi (Figura 3). Hem de tenir en compte però, que aquest no és un model determinista, i que més que l'ordre concret d'aparició de les diferents mutacions és l'acumulació de les mateixes el que condueixen a la malignitat (Fearon and Jones 1992; Kinzler and Vogelstein 1996; Toribara and Sleisenger 1995).

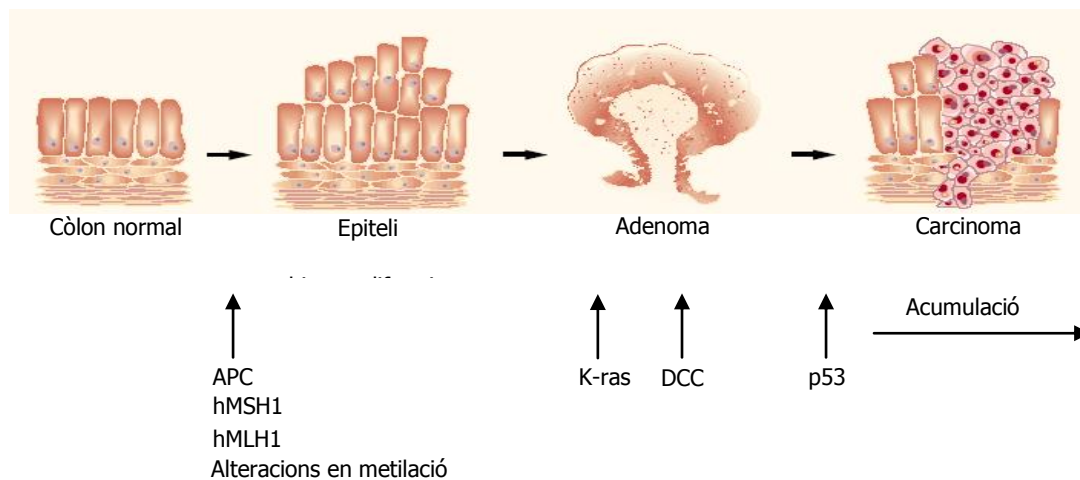


Figura 3. Model de progressió del càncer colorectal. Aquest model de progressió de càncer colorectal esporàdic basat en dades moleculars recull els gens més freqüentment inactivats en l'ordre en què preferentment s'inactiven. Adaptat de (Toribara and Sleisenger 1995).

Aquest model, tot i que incomplet, ens ajuda a entendre un fenomen d'una complexitat potser només superada per la bibliografia existent. La realitat és que s'han descrit més de 100 tipus diferents de càncer, els quals poden ser subdividits en molts casos en diferents subtipus. A aquesta complexitat en l'origen del càncer hi hem de sumar la complexitat en les seves causes moleculars. Tot plegat ha dut a alguns autors a preguntar-se per allò que és comú i fonamental en els diferents tipus de càncer, en un intent d'integrar tota la informació de la què disposem. Douglas Hanahan i Robert Weinberg han postulat que l'immens catàleg d'alteracions que trobem en càncer són el reflex de l'alteració de sis funcions essencials en la fisiologia cel·lular, que dictaminen el comportament maligne: autosuficiència en senyals de creixement, insensibilitat a senyals antiproliferatius, potencial replicatiu il·limitat, evasió de l'apoptosi, angiogènesi sostinguda i capacitat invasiva i metastàsica (Figura 4A) (Revisat en (Hanahan and Weinberg 2000)).







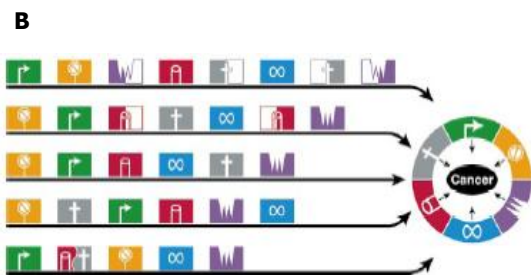
A	Capacitat adquirida	Exemple de mecanisme
	Autosuficiència senyals creixement	Activació H-Ras
	Insensibilitat inhibició creixement	Pèrdua Rb
	Evasió apoptosi	Producció IGF
	Potencial replicatiu il·limitat	Activació telomerasa
	Capacitat angiogènica	Producció inductor VEGF
	Capacitat invasiva i metastàsica	Inactivació E-cadherina

Figura 4. Les 6 capacitats adquirides de les cèl·lules tumorals. La gran quantitat d'alteracions moleculars en càncer poden ser incloses en 6 categories diferents, les quals corresponen a 6 capacitats adquirides per mitjà de guanys i pèrdues de funcions cel·lulars (A). Diferents ordres en l'adquisició d'aquestes 6 capacitats condueixen a una cèl·lula normal cap a la tumorigènesi (B). Modificat de (Hanahan and Weinberg 2000).



Tal i com ja havíem vist en el model de progressió tumoral de Vogelstein (Vogelstein et al. 1988), el model d'adquisició de noves capacitats proposat per Hanahan i Weinberg no és determinista (Figura 4B) i ens permet integrar en unes poques regles lògiques el comportament de les cèl·lules tumorals.

2. Inestabilitat genètica

El fet que el càncer pot ser considerat un procés d'evolució somàtica va ser introduït per Boveri a principis del segle XX (Boveri 1929). Sota aquesta premissa, trobem en càncer les mateixes dues forces que modelen les poblacions naturals: la variació genètica i la selecció natural, amb la diferència que aquestes dues forces s'apliquen en càncer sobre poblacions de cèl·lules i no de individus. En evolució clàssica és la mutació la que genera variació genètica en una població, sobre la qual la selecció natural proporciona una major supervivència a aquells fenotips/genotips millor adaptats a l'entorn. En càncer però, és una, o varies mutacions inicials les que proporcionen un determinat avantatge selectiu a una cèl·lula, fet que li permet expandir-se sobre les seves veïnes. De forma successiva, més mutacions permeten noves expansions clonals les quals originen una altra vegada variació genètica. Aquesta és la teoria de la selecció clonal proposada per Nowell fa més de tres dècades (Nowell 1976), la qual confirmen les dades moleculars en què es basa el model de progressió que ja hem vist (Figures 3 i 4).

En aquest context, alguns autors han suggerit que les taxes de mutació normals són insuficients per a explicar l'elevada variació genètica que observem en tumors i que per tant, les cèl·lules tumorals requereixen de l'adquisició d'algun mecanisme de inestabilitat genètica (Loeb 1991). Per contra, també existeixen evidències que demostren que l'expansió clonal, conjuntament amb la selecció natural darwiniana serien suficients per explicar aquest elevat nombre d'alteracions (Revisat en (Sieber et al. 2005; Sieber et al. 2003)).

Si bé el pes que cada un d'aquests dos mecanismes pugui tenir en la generació de variabilitat genètica en tumors pot ser objecte de discussió, el fet que existeixen tumors amb una inestabilitat genètica inherent queda fora de tot dubte, sustentat per estudis de síndromes hereditaris i també de càncers esporàdics.

L'existència d'un fenotip mutador postulat per Loeb a principis dels anys 90 (Loeb 1991) assumeix que la inestabilitat genètica s'origina per l'alteració dels mecanismes involucrats en processos com la reparació del DNA i la segregació cromosòmica. Aquestes alteracions, si bé no conferirien un avantatge selectiu *per se*, permetrien l'adquisició, a una taxa més elevada, de múltiples alteracions en altres gens rellevants per al procés tumoral (Revisat en (Sieber et al. 2003)). Estudis posteriors en CCR, tant esporàdic com síndromes hereditaris, varen donar lloc al descobriment del mecanisme fins ara més ben estudiat d'inestabilitat genètica, la inestabilitat de microsatèl·lits. Més endavant va ser també postulada l'existència de mecanismes d'inestabilitat cromosòmica.

És important fer notar que, si bé la generació de variabilitat genètica és absolutament necessària per a la superació de les barreres selectives que imposa el tumor, tant l'absència o l'excés de variabilitat (com a conseqüència d'absència de inestabilitat o una inestabilitat excessiva) poden provocar la incapacitat de les cèl·lules tumorals per adaptar-se (Figura 5) (Cahill et al. 1999).

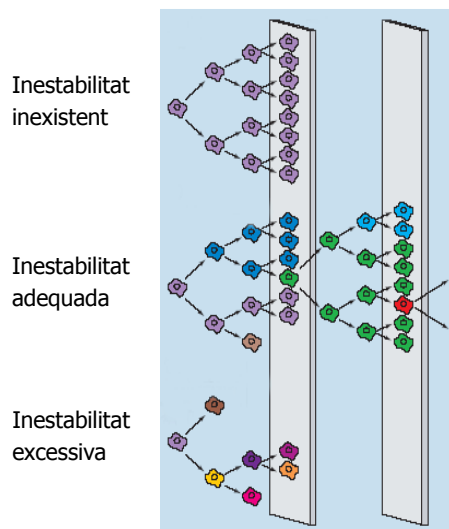


Figura 5. La variabilitat genètica és el motor de la tumorigènesi. Si les cèl·lules no són capaces de generar una població genèticament heterogènica no seran capaces d'adaptar-se a les noves condicions imposades per les barreres selectives (rectangles). Les conseqüències poden ser idèntiques si les cèl·lules són massa inestables, fet que compromet la seva viabilitat. La situació ideal és aquella en la que hi ha una inestabilitat adequada, la qual origina suficient heterogeneïtat per anar superant les diverses barreres selectives que s'imposen en l'entorn del tumor, amb les posteriors expansions clonals. Adaptat de (Cahill et al. 1999).

Inestabilitat de microsatèl·lits (*Microsatellite Instability* o **MIN**)

Aquest tipus d'instabilitat té la seva base molecular en alteracions dels mecanismes de reparació de bases mal aparellades o *mismatch* (Revisat en (Kinzler and Vogelstein 1996)). Es caracteritza per l'acumulació de milers d'alteracions somàtiques en la seqüència de DNA en forma de petites insercions, delecions i mutacions i va ser detectat quasi de forma simultània tant en pacients del síndrome hereditari HNPCC (*Hereditary Non Polyposis Colorectal Cancer*) (Aaltonen et al. 1993) i en un subgrup de tumors colorectals esporàdics (Ionov et al. 1993; Thibodeau et al. 1993). Ràpidament es va relacionar aquest tipus d'instabilitat amb defectes en la reparació i replicació del DNA, hipòtesi que va poder ser confirmada gràcies a estudis duts a terme en bacterians i de forma més important, en llevats, on mutacions en gens homòlegs als de reparació d'aparellaments incorrectes bacterians conduïen a la mateixa instabilitat observada en tumors hereditaris i esporàdics (Strand et al. 1993). Tot i que aquest tipus d'instabilitat té efectes al llarg de tot el genoma, és en les seqüències microsatèl·lit, per la seva naturalesa repetitiva, on és més fàcilment observable, fet pel qual aquest tipus d'instabilitat s'anomena de microsatèl·lits.

La instabilitat de microsatèl·lits, en el cas de tumors esporàdics, és dóna de forma molt inicial i persisteix durant el procés tumoral (Shibata et al. 1994), possibilitant l'adquisició de posteriors mutacions en oncogens i gens supressors tumorals que permetran l'avenç del tumor. L'existència d'aquest tipus d'instabilitat prediu que gens amb microsatèl·lits en regions codificants haurien de ser més propensos a patir mutacions en tumors amb aquest tipus d'instabilitat. L'any 1995 va ser descobert el primer gen, el receptor tipus II de TGF β , amb un microsatèl·lit mutat en la seva regió codificant, tant en línies de CCR (Markowitz et al. 1995) com en tumors de còlon esporàdics (Parsons et al. 1995). Altres gens que han estat descrits amb alteracions similars són IGF2R, MSH3, BAX i TCF4 (Woerner et al. 2005).

Inestabilitat cromosòmica (*Cromosomal Instability* o CIN)

Una de les principals característiques dels tumors és el gran nombre d'alteracions cromosòmiques que contenen, fet que ja va observar Boveri a principis de segle (Boveri 1929). Aquesta situació ha dut a alguns autors a postular l'existència d'algun tipus d'inestabilitat cromosòmica.

Quan parlem d'alteracions cromosòmiques en tumors, aquestes són bàsicament de dos tipus: numèriques i estructurals.

Les alteracions numèriques: es solen caracteritzar per pèrdues (aneuploidies) o guanys (poliploidies) de cromosomes sencers. Tot i que no es coneixen bé les bases moleculars que condueixen a les cèl·lules tumorals a perdre o guanyar cromosomes sencers, de bon principi es va postular que podien estar relacionades amb alteracions en els mecanismes que controlen la correcta segregació dels cromosomes durant la mitosi. Tot i que alguns *screenings* en llevat indiquen que hi ha més d'un centenar gens que poden tenir un efecte en l'estabilitat del nombre de cromosomes (Spencer et al. 1990), hi ha pocs exemples *in vivo* amb defectes en aquesta categoria de gens. Podem parlar, per exemple, dels gens de la família MAD i BUB, els quals participen directament en els *checkpoints* de mitosi que s'encarreguen de que hi hagi una correcta segregació dels cromosomes (Revisat en (Jallepalli and Lengauer 2001)).

Alteracions estructurals: sota aquesta etiqueta s'hi engloben les delecions, insercions, amplificacions, inversions i translocacions. Tot i que s'han descrit múltiples mecanismes que poden donar lloc a aquesta mena d'alteracions estructurals, els quals no esmentarem, en els últims anys han sorgit una sèrie de treballs que relacionen de forma molt important dany cromosòmic estructural i alteracions epigenètiques en tumors, punt que veurem més extensament en l'apartat d'epigenètica.

Ja per finalitzar aquest apartat, si bé hem dit que els tumors requereixen de mecanismes de producció de variació genètica per a poder anar superant les diferents barreres selectives que li són imposades, adquirint en alguns casos algun tipus d'inestabilitat genètica, també hem vist que un excés d'aquesta inestabilitat pot comportar la inviabilitat de les cèl·lules del tumor (figura 5). En aquest sentit, s'ha pogut comprovar que no es solen donar els dos tipus de inestabilitat (MIN i CIN) en un mateix tumor, existint per tant una relació inversa molt clara. Com a tota regla, es clar, s'hi ha trobat unes poques excepcions en les què es donen els dos tipus d'inestabilitat simultàniament, tant en línies cel·lulars (Morales et al. 2005) com en tumors colorectals (Goel et al. 2003).

3. Epigenètica

En biologia molecular, com en la ciència en general, cada moment històric es caracteritza per l'existència de moviments o corrents de pensament dominant que fan que determinats àrees del coneixement es desenvolupin amb una major intensitat que d'altres, moltes vegades lligats als avenços de la tècnica. Així doncs, si bé l'estudi de la genètica del càncer va dominar la dècada dels 80, podem afirmar sense cap mena de dubte que l'epigenètica ha predominat, i predomina des de finals de la dècada dels 90. Si tenim en compte el període de temps comprès entre els anys 1990 i 2006, veurem que el nombre de publicacions indexades a la base de dades PubMed (www.ncbi.nih.gov/pubmed) que contenen el terme de cerca "epigenetic" s'ha multiplicat per més de 20. El major increment en el nombre d'articles s'ha donat des de l'any 2000 ençà, gràcies en bona part als estudis realitzats en càncer (figura 6).

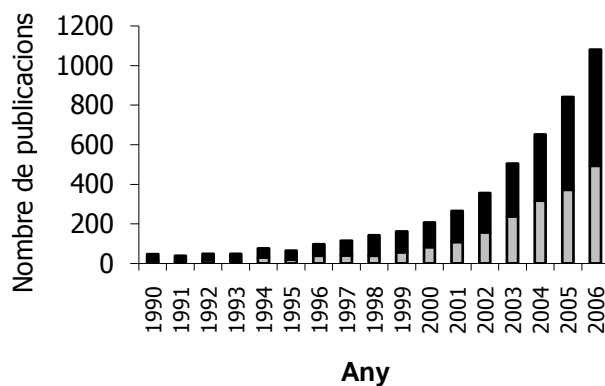


Figura 6. Evolució en el nombre de publicacions sobre epigenètica en els darrers anys. Ha estat a partir de finals dels anys 90 que la recerca en el camp de l'epigenètica s'ha expandit de forma considerable, en bona part gràcies a les publicacions referents a epigenètica en càncer (porció gris de les barres).

Com veurem en aquest apartat, tot i la gran complexitat de la literatura actual referent a epigenètica, podem dir que comencem a entendre un mecanisme de regulació les implicacions del qual tenen un fort impacte sobre les funcions cel·lulars més importants i bàsiques com són la regulació transcripcional, la replicació i el manteniment de la integritat genòmica.

Què és l'epigenètica?

Com hem vist anteriorment, el coneixement de l'origen de les paraules ens aclareix en molts casos el seu significat; en aquest cas epigenètica vol dir literalment "sobre la genètica". Si pensem en genètica i epigenètica com a dos mecanismes de regulació a mode de capes, no podríem dir inequívocament si l'epigenètica és troba realment per sobre o per sota de la genètica (com veurem més endavant, hi ha exemples per tot), però el que queda fora de tot dubte és l'existència d'una important i profunda interrelació entre ambdues. Una definició clàssica d'epigenètica és l'estudi d'aquells canvis heretables a través de la mitosi i la meiosi que afecten a l'expressió dels gens sense alterar-ne la seqüència de DNA.

La cromatina és el substrat fonamental de la regulació epigenètica

Per entendre quins són els components que juguen un paper en la regulació epigenètica hem de saber, en primer lloc, que el DNA no es troba en forma de macromolècula aïllada en el nucli, sinó formant un complex nucleoproteic que coneixem com a cromatina. Per tant, totes aquelles

reaccions metabòliques que tenen lloc al voltant del material genètic es troben cromatina, i no DNA nu, com a substrat (Revisat en (Felsenfeld and Groudine 2003)). Donada la llargada de la molècula de DNA i les reduïdes dimensions del nucli, la primera i més evident funció que es va associar a la cromatina és la de compactació del material genètic, tant en interfase com en la divisió cel·lular (Figura 7). Tot i això ja des de bon principi es va hipotetitzar que, a més a més de compactar el material genètic en el nucli, la cromatina podia jugar un paper en la regulació transcripcional.

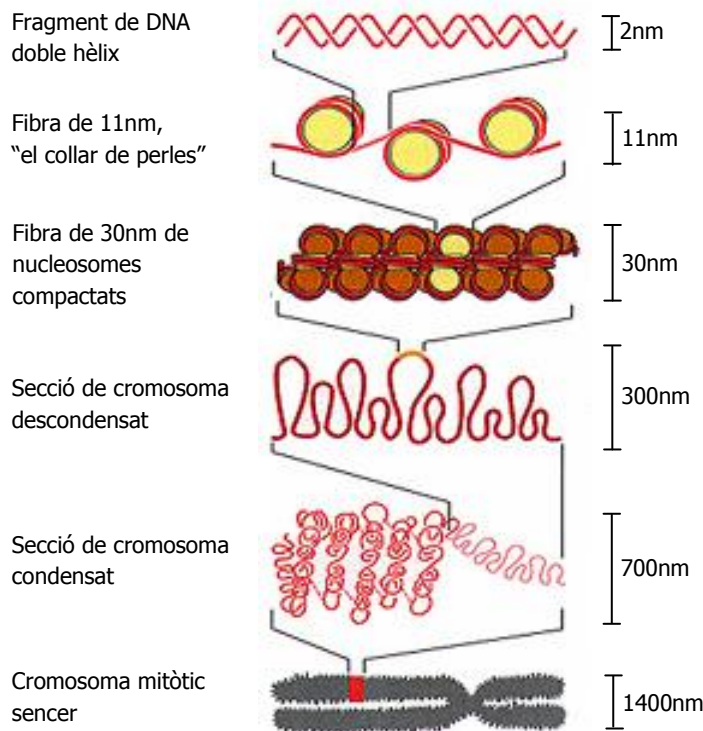


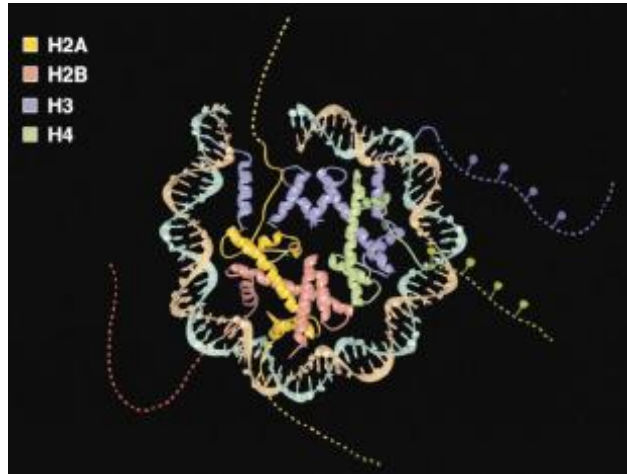
Figura 7. Del DNA al cromosoma metafàsic. La compactació de la fibra de DNA en el nucli, tant en interfase com en metafase, és una de les funcions bàsiques de les histones, les quals, conjuntament amb el DNA conformen la cromatina. L'associació de les histones amb el DNA (fibra de 11nm) comporta una compactació de l'ordre de 5 vegades. Aquesta compactació arriba fins al cromosoma descondensat interfàsic o el cromosoma condensat mitòtic, el qual representa el major ordre de compactació en cèl·lules somàtiques. Figura adaptada de (Felsenfeld and Groudine 2003).

El coneixement que la cromatina és una estructura repetitiva, formada per unitats fonamentals anomenades nucleosomes, el devem als estudis realitzats per Roger D. Kornberg i col·laboradors (Kornberg and Thomas 1974). Així doncs, el nucleosoma està format per aproximadament 165 pb de DNA, els quals donen dues voltes al voltant de l'octàmer d'histones. L'octàmer central està format per 2 subunitats de cadascuna de les histones H3, H4, H2A i H2B (Figura 8). Els nucleosomes es troben separats entre ells per fragments curts d'uns 50-80 pb de DNA (Revisat en (Kornberg and Lorch 1999)).

La regulació epigenètica té doncs dos substrats sobre els quals actuar: el DNA i les histones que l'empaqueten, és a dir, la cromatina. En els següents apartats passarem a introduir quines són les principals modificacions que afecten a ambdós components de la cromatina i quins efectes tenen sobre diferents funcions cel·lulars.

Figura 8. Estructura del nucleosoma.

La imatge mostra l'octàmer d'histones H3 (lila), H4 (verd), H2A (groc) i H2B (vermell). En el model, s'hi han afegit les cues aminoacídiques sense estructura concreta (i per tant absents en l'estructura de rajos X) que surten de l'octàmer central, el qual és bàsicament globular. Tot i que el DNA dóna dues voltes al voltant de l'octàmer, la imatge només mostra una de les dues voltes, corresponents a 73pb de DNA de doble hèlix. Model realitzat a partir de dades cristal·logràfiques. Extret de (Kornberg and Lorch 1999).

**La metilació del DNA**

L'afirmació que el material genètic està compost per polímers dels 4 nucleòtids adenina, timina, guanina i citosina, enllaçats en forma de doble hèlix, tot i que certa, és en ocasions incompleta ja que des de l'any 1948 sabem que en certs organismes existeix el que podem considerar una cinquena base nucleotídica, la 5 metil citosina (5mC) (Figura 9). Diferents enzims amb activitat metiltransferasa són capaços de catalitzar la transferència d'un grup metil des del donador SAM (s-adenosil-metionina) cap a la posició 5 prima de la citosina. En vertebrats, la metilació de la citosina es dóna bàsicament en el context de la seqüència 5'-CG-3', coneguda com a dinucleòtid CpG (Revisat en (Bird 2002; Goll and Bestor 2005)).

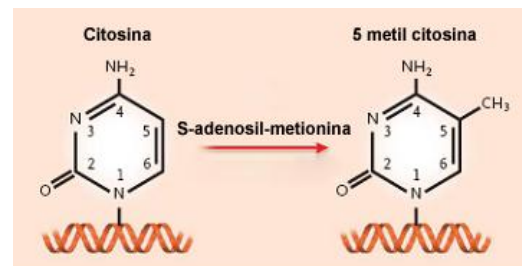


Figura 9. Estructura de la citosina i la 5 metil citosina. Com el seu nom indica, la 5 metil citosina conté un grup metil en posició 5 prima de l'anell aromàtic, el qual és addicionat per enzims amb activitat metil transferasa a partir de la s-adenosil-metionina. Modificat de (Herman and Baylin 2003).

La metilació del DNA en cèl·lules normals

Per a poder entendre quin paper juga la metilació en un context patològic com és el CCR, primer hem de conèixer quin o quins papers juga la metilació del DNA en un context fisiològic normal. Els genomes de mamífers, i els de vertebrats en general, es caracteritzen per tenir els nivells més alts de 5mC de tot el regne animal, en comparació amb els genomes de invertebrats, els quals posseeixen nivells més baixos o inexistents (Taula 1) ((Tweedie et al. 1997); Revisat en (Bird 2002)). L'ús d'enzims de restricció sensibles i insensibles a la metilació va permetre, a principis de la dècada dels 80, començar a respondre dues preguntes bàsiques relacionades amb la metilació del DNA en mamífers i altres vertebrats: quins són els nivells globals de metilació i com es distribueixen al llarg del genoma?

Taula 1. Contingut de 5mC en diversos organismes model.

Organisme	Classificació	Patró de metilació ^a
<i>Saccharomyces cerevisiae</i>	Organisme eucariota unicel·lular	Inexistent
<i>Caenorhabditis elegans</i>	Invertebrat nematode	Inexistent
<i>Drosophila melanogaster</i>	Invertebrat insecte	Inexistent ¹
<i>Apis mellifera</i>	Invertebrat insecte	Fraccionat ²
<i>Ciona intestinalis</i>	Invertebrat cordat	Fraccionat
<i>Homo sapiens</i>	Vertebrat	Global
<i>Arabidopsis thaliana</i>	Vegetal	Global

a. D'acord amb la classificació realitzada per (Tweedie et al. 1997), basada en la sensibilitat a enzims de restricció.

1. Bioquímicament s'hi han detectat nivells molt baixos de 5mC, sobretot en el context TpG (Gowher et al. 2000; Lyko et al. 2000).

2. S'hi han detectat nivells baixos de 5mC, en el context de CpG (Wang et al. 2006).

Treballant amb DNA de diferents teixits humans, Melanie Ehrlich i col·laboradors varen estimar que al voltant del 1% del genoma humà està compost per 5mC (Ehrlich et al. 1982), o el que és el mateix, que el genoma humà conté al voltant de 30 milions de citosines metilades. Estudis posteriors han situat aquests percentatges entre el 3 i el 5% del total de bases (Revisat en (Bird 2002; Ehrlich 2002)). Del total de citosines metilades, entre el 70 i el 80% es troben en posició 5' a una guanina, es a dir, dins del dinucleòtid CpG (Revisat en (Bird 2002)). Després de més de 20 anys de recerca en aquest camp sabem que aquesta quantitat de 5mC no es distribueix a l'atzar, fet que està directament relacionat l'especial distribució dels dinucleòtids CpG en el genoma de mamífers.

Els dinucleòtids CpG es distribueixen heterogèniament: les illes CpG

Tot i que les primeres reaccions de seqüenciació no eren tant ràpides ni senzilles com les que avui en dia tenim a l'abast de forma rutinària, varen permetre dur a terme els primers anàlisis de seqüències curtes. Un dels resultats més sorprenents va ser veure com, en genomes de vertebrats, el dinucleòtid CpG es troba a una freqüència 2/3 inferior a la que caldria esperar per la composició de bases (Josse et al. 1961; Swartz et al. 1962). La ja no tant recent publicació de la seqüència del genoma humà (Lander et al. 2001) ha permès realitzar estudis globals sobre la freqüència i distribució d'aquest i altres dinucleòtids, els quals han confirmat, a gran escala, les dades que coneixíem des de principis dels anys 60 (Lander et al. 2001; Venter et al. 2001).

Dos fenòmens expliquen aquesta manca de dinucleòtids CpG en genomes de mamífer: en primer lloc sabem que tant la citosina com la 5mC tenen tendència a patir desaminació hidrolítica espontània, essent aquesta taxa 2.2 vegades superior en el cas de la 5mC en DNA de cadena doble (Shen et al. 1994). En segon lloc, la desaminació de la citosina i de la 5mC dona lloc a uracil i timina, respectivament. Donat que la presència de uracil és estranya en un context de DNA, aquesta és reparada eficientment (Lindahl 1993). En canvi, el *mismatch* provocat per

la desaminació a timina pot ser resolt incorrectament donant lloc a una mutació de tipus transició, amb el corresponent increment dels dinucleòtids TpG i CpA (Bird 1980). Així doncs, una major tendència de la 5mC a desaminar espontàniament a timina i una reparació més ineficient dels aparellaments incorrectes creats per aquesta desaminació expliquen la baixa freqüència del dinucleòtid CpG en genomes amb patrons de metilació globals del DNA.

Tot i aquesta disminució global en el nombre de dinucleòtids CpG en el genoma de mamífers, existeixen regions curtes de DNA (de centenars a pocs milers de parells de bases) en les quals la freqüència de CpGs és més elevada, aproximant-se als valors que esperaríem per atzar d'acord amb la seqüència de bases. Són les anomenades illes CpG (Figura 10), les quals estan associades a les regions promotores del 76% dels gens del genoma humà, tant *house-keeping* com d'expressió dependent de teixit (Revisat en (Goll and Bestor 2005)). Si bé la primera definició d'illa CpG la devem a Gardiner-Garden i col·laboradors (Gardiner-Garden and Frommer 1987), revisions posteriors basades en dades de seqüència de cromosomes complets han

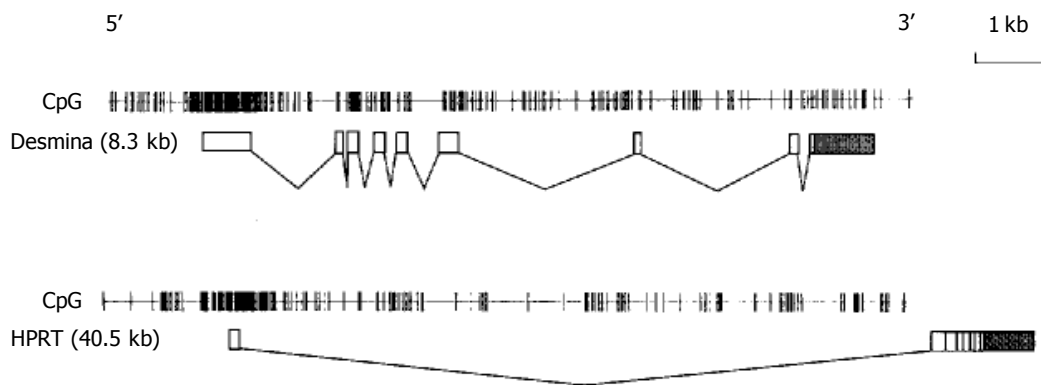


Figura 10. Les illes CpG. En la figura es poden veure 2 gens (Desmina i Hipoxantina fosforibosil transferasa o HPRT) amb els exons (caixes) enllaçats per els introns (línies). Els dinucleòtids CpG (línies verticals que travessen línia horitzontal) estan clarament infrarepresentats en ambdós gens excepte al voltant de les regions 5' promotores, que en el cas d'aquests dos gens estan al voltant del primer exó. Aquesta major densitat de CpGs configura les illes d'ambdós gens en els seus promotors. Modificat de (Cross and Bird 1995).

permès establir una sèrie de criteris per a definir les illes CpG d'una forma més estricta (Takai and Jones 2002). Així doncs, una seqüència genòmica, per tal de ser considerada illa CpG, ha de complir els següents tres criteris:

1. La seva longitud ha de ser igual o superior a 500 pb
2. El seu contingut de G + C ha de ser igual o superior a 0.55
3. El quocient entre el nombre de dinucleòtids CpG observats i els esperats ha de ser igual o superior a 0.6

Tot i que en aquesta tesi ens centrarem bàsicament en les illes CpG associades a les regions promotores (aquelles situades en l'extrem 5' del gen), també existeixen illes CpG intragèniques i associades als extrems 3' de diferents gens.

Patrons de metilació: com s'estableixen?

Una de les propietats dels patrons de metilació de cèl·lules somàtiques, és la seva estabilitat (Revisat en (Bird 2002)). Això no vol dir però, que els patrons de metilació de mamífers siguin estables al llarg del desenvolupament. Excloent els canvis locals de metilació associats a la inactivació del cromosoma X, els gens imprimats i els gens silenciats específics de teixit (els quals es veurem més endavant), els nivells de metilació sofreixen canvis molt importants en un moment molt determinat del desenvolupament en mamífers. Concretament, unes hores després de la fertilització, els patrons de metilació aportats pel genoma del progenitor masculí són esborrats per mitjà d'un procés que es creu actiu (per l'absència de replicació del DNA en aquest punt) (Mayer et al. 2000), mentre que el genoma aportat pel progenitor femení es desmetila de forma passiva en les subsegüents divisions cel·lulars per manca de activitat metiltransferasa de manteniment (Revisat en (Li and Bird 2007)), arribant-se a perdre fins al 50% de la metilació global. Després d'aquest procés de desmetilació global, els patrons de metilació que es mantindran durant la vida adulta s'estableixen just després de la implantació de l'embrió, en el moment de la gastrulació (Revisat en (Bird 2002; Li and Bird 2007)).

Com són els patrons de metilació en mamífers?

Ja hem dit que la metilació del DNA afecta bàsicament al dinucleòtid CpG i que aquest es distribueix d'una forma molt especial en el genoma, fet que condiciona la distribució de 5mC en el genoma de mamífers. Aquests posseeixen el que s'anomenen patrons de metilació globals (Taula 1), és a dir, aquells en els que una fracció majoritària del genoma està fortament metilada, mentre que la fracció minoritària restant es manté bàsicament lliure de metilació (Tweedie et al. 1997; Yoder et al. 1997)(Revisat en (Bird 2002)).

Són molts els treballs que han concentrat els esforços en estudiar quin tipus de seqüències trobem en cada compartiment de metilació. Inicialment, aquests estudis varen ser realitzats per mitjà de l'ús d'enzims de restricció, els quals només permeten "veure" l'estat de metilació d'una o unes poques citosines. El que coneixem de forma inequívoca és que aquesta fracció majoritària metilada correspon en gran mesura a seqüències repetides diverses (disperses i en tàndem), seqüències intergèniques i seqüències intragèniques, en resum, tot allò que queda fora del conjunt de seqüències amb una funció reguladora. Estudis molt recents, realitzats per mitjà de seqüenciacions massives de regions metilades i no metilades, han permès un anàlisi més detallat de quines són les característiques de les seqüències que conformen cada compartiment i quin és el seu grau de metilació. Aquest estudi ha confirmat la presència majoritària de DNA repetitiu i regions no reguladores en la fracció metilada i illes CpG, regions promotores i primers exons en la fracció no metilada (Rollins et al. 2006).

Aquest patró tan especial de metilació respon a una de les funcions més bàsiques de la metilació del DNA: el silenciament estable de les seqüències metilades. La meitat del genoma humà està composta per seqüències repetides (Lander et al. 2001), les quals, com acabem de veure, acumulen la major part de la metilació en mamífers (Revisat en (Yoder et al. 1997)). Aquest fet ha dut a alguns autors a suggerir que la metilació d'aquests elements és un

mecanisme de supressió de la seva capacitat de transposar-se i per tant, de posar en perill la integritat del genoma (Yoder et al. 1997). Per contra, l'estat bàsicament no metilat d'una fracció majoritària de les illes CpG i altres elements reguladors està associat a un estat transcripcional actiu (o la potencialitat per expressar-se) dels gens a què estan associats. Aquest estat lliure de metilació de les illes CpG explica el perquè aquestes regions del genoma tenen un contingut tant elevat de dinucleòtids CpG, a més a més dels possibles constrenyiments selectius que puguin operar sobre aquestes seqüències per tal de mantenir la seva funció.

Des d'un punt de vista molt simplista, el genoma humà se'ns presenta com a un gran desert de dinucleòtids CpG fortament metilats, entre el qual hi ha disperses illes riques en CpGs, lliures de metilació.

Les activitats DNA metiltransferasa

Com ja hem vist, la metilació de la citosina es dona bàsicament en el context del dinucleòtid CpG (Revisat en (Bird 2002)). El fet que aquest dinucleòtid és un palíndrom, tot i que el més petit imaginable, va fer sorgir la idea, de forma independent en dos laboratoris, que aquest podria ser un sistema de memòria cel·lular (Holliday and Pugh 1975; Riggs 1975). Investigacions posteriors sobre els patrons de metilació de gens ribosòmics en *Xenopus laevis* varen demostrar que la metilació (o l'absència de metilació) es donava de forma simètrica en el dinucleòtid CpG d'ambdues cadenes de DNA, fet que suggeria que els patrons de metilació podrien efectivament ésser copiats a través de la replicació del DNA (Bird 1978). Aquestes observacions varen predir l'existència de dues activitats enzimàtiques clau: una capaç d'establir *de novo* els patrons de metilació i una altra capaç de copiar-los i mantenir-los, de forma semiconservativa, a través dels processos de divisió cel·lular.

El primer enzim amb activitat metiltransferasa aïllat (l'únic aïllat bioquímicament) es va anomenar Dnmt1 (Bestor and Ingram 1983) i, tot i que presentava certa activitat enfront de DNA no metilat, la seva preferència per DNA hemimetilat va fer pensar ràpidament que aquesta era l'activitat de manteniment que s'estava cercant. Estudis posteriors varen confirmar que Dnmt1 s'associa amb les forques de replicació durant la fase S (Leonhardt et al. 1992), fet que sustenta el paper de la Dnmt1 com a metiltransferasa de manteniment. Tot i això, estudis recents han demostrat que aquesta associació es pot estendre durant la fase G2 i M (Easwaran et al. 2004). La inactivació selectiva en homozigosi d'aquest enzim en cèl·lules mare embrionàries (ESC) comporta una reducció dels nivells de metilació en un 65% que no comprometen la viabilitat cel·lular. El mateix experiment dut a terme en embrions de ratolí comporta una reducció similar en els nivells de metilació, que en aquest cas comporten la mort de l'embrió abans del naixement (Li et al. 1992). Aquests resultats suggereixen que la metilació juga un important paper en l'establiment dels patrons de metilació durant el desenvolupament, però no són essencials per a la viabilitat de cèl·lules mare embrionàries (Revisat en (Li and Bird 2007)).

El descobriment de l'enzim Dnmt1 va permetre reconèixer els motius de DNA específics que conferien activitat metiltransferasa. Això va possibilitar la cerca bioinformàtica en les bases de

dades d'EST d'altres proteïnes que poguessin contenir aquests dominis i per tant, tenir activitat metiltransferasa. D'aquesta manera es varen descobrir la Dnmt2, la qual té una activitat metiltransferasa mínima (Okano et al. 1998b) i les Dnmt3a i Dnmt3b (Okano et al. 1998a). Tant la Dnmt3a com la 3b no mostraven preferència per DNA hemimetilat i la seva disrupció en ESC impedia a aquestes cèl·lules metilar *de novo* provirus i elements repetitius (Okano et al. 1999) fet que va suggerir que aquestes dues DNA metiltransferases eren efectivament els enzims que proporcionen la metilació *de novo*. Per a una extensa i excel·lent revisió de les activitats metiltransferasa, no només en mamífers, sinó en plantes i invertebrats, referir-se a (Goll and Bestor 2005).

Quines són les funcions de la metilació del DNA?

Al llarg dels últims 25 anys s'han dut a terme un gran nombre d'experiments destinats a desxifrar quin és el paper de la metilació del DNA. Quasi tots ells arriben a la mateixa conclusió: la metilació del DNA s'associa al silenciament transcripcional. Com més coneixement adquirim sobre aquest fenomen, més ens adonem però que el silenciament transcripcional és només una de les múltiples cares d'un fenomen molt més ampli: el paper de la metilació del DNA en la regulació de l'arquitectura de la cromatina.

Tot i que com abordarem més endavant en aquesta introducció, un dels papers més importants de la metilació del DNA és a nivell genòmic en la supressió de l'activitat dels elements transposables, en aquest apartat ens centrarem exclusivament en els efectes a nivell gènic. En aquest context, una de les funcions de la metilació del DNA més llargament hipotetitzades és el seu possible paper en la regulació transcripcional de gens durant el desenvolupament, suggerint l'existència de gens metilats (i per tant silenciats) que podrien desmetilar-se i per tant reactivar-se en moments clau del desenvolupament. Fins a la data però, no es coneix cap gen que de forma convincent mostri aquest comportament en processos de desenvolupament ((Walsh and Bestor 1999); Revisat en (Goll and Bestor 2005)). Les tres principals situacions fisiològiques normals en les que la metilació del DNA sí que juga un important paper com a mecanisme de silenciament són la inactivació del cromosoma X en mamífers, l'establiment de l'imprinting i el silenciament específic de teixit.

Inactivació del cromosoma X en mamífers

Els genomes d'eucariotes superiors són sistemes d'expressió gènica molt finament regulats, els quals compensen la diferent dotació de cromosomes sexuals que hi ha entre els dos sexes ajustant els nivells d'expressió dels gens en aquests cromosomes. En dones, el fet de tenir dos cromosomes X planteja el problema de tenir una major dosi gènica que els homes per els mateixos gens, situació que ha estat resolta evolutivament silenciant un dels dos cromosomes X per mitjà d'una cascada d'esdeveniments iniciats per el gen Xist, que inclouen de forma seqüencial, el silenciament transcripcional, la reestructuració de la cromatina i finalment la metilació del DNA (Revisat en (Heard 2004); tot i que la metilació del DNA no és el seu punt d'atenció, aquesta revisió recopila el conjunt d'esdeveniments que condueixen a la inactivació del cromosoma X en mamífers). Estudis posteriors han demostrat que si bé les illes CpG del

cromosoma inactiu estan més metilades que les del cromosoma X actiu, aquest últim presenta nivells de metilació intragènica més alts (Hellman and Chess 2007). Aquests resultats suggereixen que no són els nivells de metilació sinó la distribució de la 5mC la que varia de forma ostensible en la inactivació del cromosoma X.

Imprinting

L'imprinting consisteix en el silenciament selectiu d'un determinat al·lel en funció de si aquest prové del progenitor masculí o femení. Per a molts dels gens imprintats s'ha descrit la presència de regions de control del imprinting (conegudes com a ICRs), l'estat de metilació de les quals determina l'estat transcripcional del gen al qual estan associades. La metilació de les ICRs es realitza durant la gametogènesi, tant masculina com femenina, per mitjà de la metil transferasa *de novo* DNMT3b (Revisat en (Delaval and Feil 2004)). Un dels casos millor estudiats és el del locus IGFR2/H1 en la regió telomèrica del braç curt del cromosoma 11. En aquest cas concret, trobem metilació del DNA sobre el ICR patern, el qual provoca el silenciament del gen H19 i l'expressió del gen IGF2. Per contra, en la dona, aquest ICR es manté desmetilat, fet que provoca que s'expressi el gen H19 i no el IGF2 (Revisat en (Delaval and Feil 2004)). Estudis directament enfocats a la recerca de illes CpG metilades en teixit normal han permès descobrir noves regions amb gens imprintats i gens metilats específicament en determinats teixits (Strichman-Almashanu et al. 2002).

Metilació del DNA específica de teixit

Els processos de desenvolupament és caracteritzen per la creació de múltiples tipus cel·lulars, amb formes i funcions clarament diferenciades. A partir d'una cèl·lula ancestral comú (el zigot) es deriven diferents llinatges cel·lulars, els quals donaran lloc, per processos de diferenciació, al repertori de cèl·lules somàtiques i germinals que conformen l'adult. Aquests processos de diferenciació és caracteritzen per la presa seqüencial de decisions que finalment configuren els patrons d'expressió específics que determinen la identitat d'un tipus cel·lular concret. Entre aquestes decisions hi ha tant l'activació com la inactivació de determinats gens. Doncs bé, entre els gens que un determinat tipus cel·lular té silenciats en un teixit específic, n'hi ha que es presenten acompanyats de metilació del DNA en les illes CpG de les seves regions promotores (Song et al. 2005). També tenim el cas contrari, es a dir, gens actius no metilats exclusivament en determinats teixits (Schilling and Rehli 2007). De totes maneres, el baix nombre de gens descoberts fins a la data que presenten aquests patrons de metilació fa pensar que aquest fenomen no és molt freqüent.

Quins mecanismes condueixen al silenciament gènic dels gens metilats?

La resposta més directa i evident a aquesta pregunta és la següent: la metilació del DNA interfereix amb la unió de factors que s'uneixen al DNA en regions riques en dinucleòtids CpG. Tot i que senzilla i elegant, aquesta resposta ha demostrat ser certa en un nombre restringit de contextos. Així doncs es coneixen alguns factors que són incapaços d'unir-se a la seva seqüència de reconeixement quan aquesta està metilada (Watt and Molloy 1988). Un dels més

extensament estudiats és el factor CTCF, el qual és essencial per al silenciament selectiu dels gens imprintats IGF2/H19 ((Bell and Felsenfeld 2000; Hark et al. 2000; Szabo et al. 2000); Revisat en (Delaval and Feil 2004)). Alguns *screenings* han permès identificar més de 200 llocs no metilats, prèviament no identificats, de unió de CTCF en el genoma de ratolí, fet que suggereix un paper d'aquesta proteïna, més enllà de les funcions d'imprinting, en el manteniment de dominis transcripcionals aïllats al llarg de tot el genoma (Mukhopadhyay et al. 2004).

La visió oposada (i més complexa) és aquella que proposa que la metilació del DNA atreu, més que no pas allunya activitats, assumint que la metilació del DNA és una senyal que pot ser reconeguda per activitats repressores de l'expressió gènica. Aquesta hipòtesi prediu l'existència de proteïnes capaces de llegir les marques de metilació. Aquesta predicció va deixar de ser-ho al descobrir-se, per mètodes bioquímics, el complex MeCP1 (Meehan et al. 1989). La primera proteïna aïllada de forma bioquímica que s'unia a CpGs metilats es va anomenar MeCP2. L'estudi d'aquesta proteïna va revelar la presència d'un domini de unió a 5mC (el *Methyl Binding Domain* o MBD), fet que va permetre realitzar cerques bioinformàtiques (com ja s'havia fet en el cas de les DNMTs) d'altres proteïnes que continguessin aquest domini. Així es va reconèixer l'existència d'una família de proteïnes de unió a 5mC, les quals varen ser anomenades *Methyl Binding Domain proteins* o MBDs (MBD1, MBD2, MBD3 i MBD4) (Revisat en (Goll and Bestor 2005; Hendrich and Tweedie 2003; Jorgensen and Bird 2002)). Són mutacions en MeCP2 les causants dels desordres provocats en pacients amb síndrome de Rett. Aquests pacients són bàsicament dones (MeCP2 està en el cromosoma X) que es caracteritzen per presentar severes defectes neurològics i motors. Si bé els patrons de metilació d'aquests pacients són normals, el mecanisme de silenciament no funciona correctament per manca de reconeixement eficaç de la marca de metilació (Bienvenu and Chelly 2006).

La clau per entendre com aquestes proteïnes de unió a 5mC poden conduir al silenciament gènic dels gens metilats va venir pel descobriment que MeCP2 s'associa, en els gens metilats, a les proteïnes amb capacitat de modificar covalentment les histones. Aquest descobriment va posar en evidència que dos vies de investigació que fins aleshores havien transcorregut bàsicament paral·leles (la metilació del DNA i les modificacions de histones), s'acabaven d'entrecruar en el camí cap a la integració dels coneixements de la regulació epigenètica de l'activitat transcripcional.

Les modificacions de les histones

Les histones que empaqueten el DNA es troben entre les proteïnes més conservades al llarg de l'escala evolutiva. De la seva estructurada part globular central, en sorgeixen cues (els extrems N-terminals), les quals no posseeixen una estructura concreta (Figura 8 i 11). Una de les característiques més interessants de les histones és la gran quantitat de modificacions posttraduccionals o PTMs (*Posttranslational Modifications*) a què poden ser sotmeses. Aquestes modificacions es donen quasi de forma exclusiva en les cues de les histones per la seva fàcil accessibilitat, i consisteixen en l'addició covalent de diferents grups químics en posicions concretes (Figura 11 i Taula 2).

Les primeres evidències que les histones, a través de les PTMs, podien jugar un paper en la regulació transcripcional van venir dels estudis realitzats en el llevat *Saccharomyces cerevisiae* a finals de la dècada dels 80. Aquest organisme model va ser, i encara és, una excel·lent plataforma d'estudi de les PTMs i la seva relació amb l'activitat transcripcional per dues raons fonamentals: la conservació del sistema de regulació (el que és cert per llevats sol ser cert per a eucariotes superiors) i l'existència de còpies úniques per a cada gen de les histones en el llevat. Aquest últim fet va permetre realitzar estudis genètics que en humans (els quals posseïm múltiples còpies i variants de cada gen) haurien estat molt més difícils, per no dir impossibles (Revisat en (Kouzarides 2007)). La taula 2 resumeix les principals modificacions descrites així com els residus que solen afectar els quals poden estar més o menys distants del domini globular de les histones (Figura 11). Aquestes poden consistir en l'addició de petits grups químics com acetilacions, metilacions i fosforilacions, o bé l'addició de grups químics més grans com són la ubiquitinització o la sumoïlització. Recentment s'han descrit també altres modificacions que involucren processos una mica diferents, com són la ADP ribosilació de glutamines, la deiminació d'arginines o la isomerització de prolines (Revisat en (Kouzarides 2007)), totes tres relacionades amb la regulació de l'activitat transcripcional.

Taula 2. Tipus de modificacions covalents de les histones.

Tipus de modificació	Paper en la transcripció	Llocs modificats
Acetilació	Activador	H3 (K9, K14, K18, K56) H4 (K5, K8, K12, K16) H2A H2B (K6, K7, K16, K17)
Metilació	Activador	H3 (K4, K36, K79)
	Repressor	H3 (K9, K27) H4 (K20)
Fosforilació	Activador	H3 (S10)
Ubiquitinització	Activador	H2B (K123)
	Repressor	H2A (K119)
Sumoïlització	Repressor	H3
		H4 (K5, K8, K12, K16)
		H2A (K126)
		H2B (K6, K7, K16, K17)

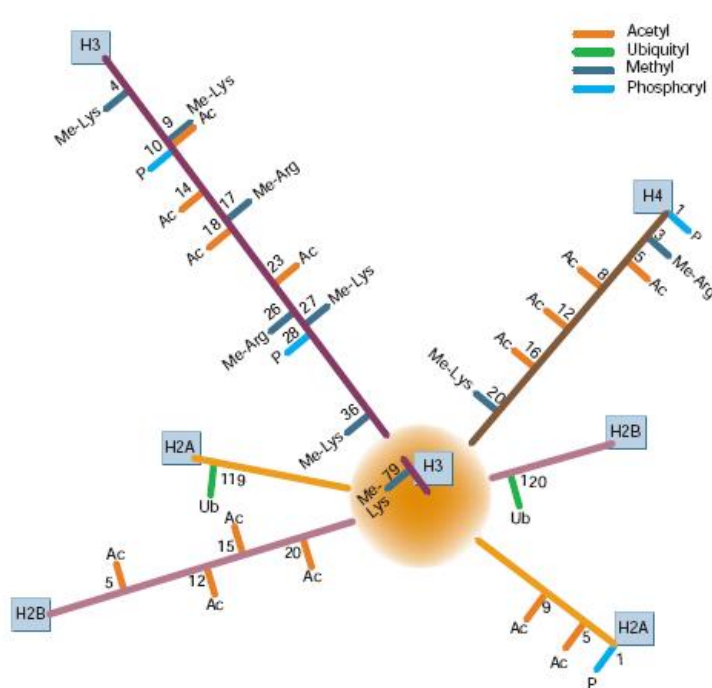


Figura 11. Les modificacions de les histones. A partir d'una estructura central globular, sorgeixen les cues de les histones, les quals poden sofrir diferents modificacions covalents en residus concrets (numerats a partir de l'extrem N-terminal). En la figura veiem acetilacions (Ac), metilacions (Me), fosforilacions (P) i ubiquitinacions (Ub). Es coneixen poques modificacions que afectin a residus en el domini globular de les histones com és la metilació de la lisina 79. Extret de (Felsenfeld and Groudine 2003).

La recerca en aquest camp s'ha vist incrementada enormement des que a principis dels anys 90 es va desenvolupar una tècnica d'anàlisi de la cromatina coneguda com a immunoprecipitació de cromatina o ChIP. Aquest mètode permet conèixer quines proteïnes s'associen a una determinada seqüència de DNA *in vivo* i ha estat instrumental en el descobriment de l'associació que existeix entre les modificacions de les histones i l'activitat transcripcional.

El gran nombre de PTMs conegudes que s'associen a diferents estats transcripcionals va dur a alguns autors a hipotetitzar l'existència d'un codi d'histones (Jenuwein and Allis 2001; Strahl and Allis 2000; Turner 2000). Segons aquesta hipòtesi, diferents combinacions de PTMs sobre les histones actuen de base de reconeixement i reclutament de diferents factors modificadors de les histones i de l'arquitectura de la cromatina. L'existència de qualsevol codi prediu dues qüestions fonamentals: la primera és l'existència d'una senyal (en aquest cas les PTMs) i la segona és l'existència de dominis proteics específics que les reconguin.

A continuació tractarem en detall les dues modificacions d'histones que són més rellevants per aquesta tesi, l'acetilació i la metilació, i els mecanismes a través dels quals aquestes modificacions tenen conseqüències sobre l'expressió gènica i l'estructura de la cromatina.

Acetilació de les histones

Aquesta va ser la primera PTM que va ser associada de forma invariable i positiva amb l'activitat transcripcional. L'addició de grups acetil es dona de forma exclusiva en residus de lisina en les cues de la histona H3, H4, H2A i H2B (Taula 2)(Revisat en (Kouzarides 2007; Kouzarides and Berger 2007)), tot i que ens centrarem exclusivament en les dues primeres.

Ja l'any 1964, treballs realitzats per Allfrey i col·laboradors varen posar de manifest que l'acetilació de les histones podia tenir un paper facilitador de l'activitat transcripcional (Allfrey et al. 1964), hipòtesi que va ser confirmada posteriorment per mitjà de diverses aproximacions bioquímiques. Una evidència més directa va arribar 25 anys més tard, per mitjà de l'ús d'anticossos contra la histona H4 acetilada en eritròcits de pollastre (Hebbes et al. 1988). Aquests experiments demostren que els gens actius en el locus de la β -globina d'eritròcits estan enriquits per aquesta marca d'acetilació, mentre que el gen de la ovoalbúmina (inactiu) no presenta aquest enriquiment. Estudis posteriors realitzats pel mateix grup indiquen però que l'associació de l'estat hiperacetilat d'aquests gens en el locus de la β -globina no està associat necessàriament amb transcripció sinó amb la seva capacitat d'activació (Hebbes et al. 1992). Usant com a model el llevat *S.cerevisiae* es va poder demostrar que l'eliminació de la cua N-terminal de la histona H4 suprimia parcialment l'expressió d'alguns gens, situació similar a la que s'aconseguia mutant residus de lisina específics (Durrin et al. 1991). Si bé aquests estudis indiquen que l'estat hiperacetilat es correlaciona amb activitat transcripcional, Jeppesen i col·laboradors varen analitzar el cas complementari. Usant immunotinció sobre cromosomes metafàsics murins varen veure que el cromosoma X inactiu està bàsicament desacetilat (excepte en aquelles petites regions on hi ha gens que escapen al silenciament), mentre que la resta de cromosomes mantenen nivells alts d'acetilació (Jeppesen and Turner 1993). Els autors suggereixen doncs que la manca d'acetilació de la histona H4 és característica tant de la heterocromatina facultativa com de la constitutiva.

Els nivells d'acetilació depenen de dues activitats: les HAT i les HDAC

Els nivells d'acetilació de les histones són, com veurem una mica més endavant, variables tant en l'espai com en el temps. Això s'aconsegueix per mitjà de l'acció de dues activitats totalment oposades: l'activitat acetiltransferasa (HAT), la qual com el seu nom indica transfereix grups acetil, i l'activitat desacetilasa (HDAC), la qual els elimina. Sorprenentment, ambdues activitats varen ser aïllades quasi de forma simultània en el ciliat *Tetrahymena* (activitat HAT)(Brownell et al. 1996) i cèl·lules de mamífer (activitat HDAC)(Taunton et al. 1996). En ambdós casos, les proteïnes aïllades presentaven una alta homologia a dues proteïnes de llevat prèviament caracteritzades per jugar un paper en la regulació transcripcional, l'activador Gcn5 en el cas de la HAT i el corepressor Rpd3p en el cas de la HDAC. D'acord amb el seu paper en l'activació transcripcional, les activitats HAT s'associen a complexos activadors de la transcripció, mentre que les activitats HDAC s'associen a complexos repressors (Revisat en (Kouzarides and Berger 2007)).

Histone acetyl transferases (HATs)

Diferents HATs mostren especificitat per l'acetilació de diferents residus de lisina en diferents histones, però en conjunt, totes les histones poden presentar residus acetilats (Taula 2). Existeixen tres principals famílies de HATs: la família relacionada amb Gcn5 coneguda com a GNAT, la família MYST i la família de CBP/p300, les quals es poden trobar contingudes en diferents complexos activadors (Taula 3)(Revisat en (Sterner and Berger 2000)). Si bé la família GNAT modifica preferentment residus en la histona H3 i la família MYST en la histona H4, la

família CBP/p300 no mostra preferència per cap de les dues (Taula 3). Membres de les tres famílies poden acetilar altres substrats que no siguin histones (Glozak et al. 2005).

Taula 3. Classificació de HATs conegudes i putatives.

HAT	Organisme	Funció	Especificitat ¹	Complex HAT
Superfamília				
GNAT				
Hat-1	Llevat > humans	Cap coneguda	H4	Llevat HAT-B, HAT-A3
Gcn-5	Llevat > humans	Coactivador	H3/H4	Llevat ADA, SAGA; humans GCN5, STAGA, TFIC
PCAF	Humans, ratolí	Coactivador	H3/H4	Humans PCAF
Elp3	Llevat	Elongació	ND	Holoenzim Pol II
Hpa2	Llevat	Desconegut	H3/H4	
Família MYST				
Sas2	Llevat	Silenciament	ND	
Sas3	Llevat	Silenciament	H3/H4/H2A	NuA3
Esa1	Llevat	Progressió cicle	H4/H3/H2A	NuA4
MOF	<i>Drosophila</i>	Compensació dosi	H4/H3/H2A	MSL
Tip60	Humans	Interacció HIV-Tat	H4/H3/H2A	Tip60
MOZ	Humans	Leucèmia	ND	
MORF	Humans	Desconeguda	H4/H3	
HBO1	Humans	Interacció ORC	ND	HBO1
CBP/p300	Varis organismes	Coactivador global	H2A/H2B/ H3/H4	
TAF _{II} 250	Llevat > humans	Associat a TBP	H3/H4	TFIID
TFIIIC		RNA pol III iniciació		TFIIIC
TFIIIC220	Humans		ND	
TFIIIC110	Humans		ND	
TFIIIC90	Humans		H3	
Coactivadors		Coactivadors de receptors		
SRC-1	Humans, ratolí	en el nucli. Regulació	H3/H4	
ACTR	Humans, ratolí	transcripcional en resposta	H3/H4	
TIF2	Humans, ratolí	a hormones	ND	

1. En negreta es mostra la histona preferentment modificada per l'enzim recombinant *in vitro*. ND: no determinat.

Histone deacetylases (HDACs)

Existeixen nombroses proteïnes amb activitat HDAC conservades des de llevats a humans, les quals han estat classificades en tres grups principals, HDAC classe I, II i III (Revisat en (Kouzarides and Berger 2007; Yang and Seto 2003)). És interessant el fet que els membres de la tercera classe, els relacionats amb la proteïna Sir2 de llevats, depenen del cofactor NAD per a la seva activitat, mentre que els membres de les classes I i II no en depenen. La taula 4 resumeix les HDAC de llevats i mamífers conegudes, així com els complexos corepressors en els quals es troben.

Taula 4. Classificació de HDACs conegudes en llevat i mamífers

Classe	<i>S. cerevisiae</i>		Mamífers	
	HDAC	Complex	HDAC	Complex
I	Rpd3	Sin3	HDAC1,-2	Sin3, NuRD, CoREST
	Hos1	Desconegut	HDAC3	N-CoR, SMRT
	Hos2	Set3	HDAC8	Desconegut
II	Hda1	Hda2/Hda3	HDAC4,-5,-7,-9	Desconegut
			HDAC6	p97/PLAP
			HDAC10	Desconegut
III	Sir2	Sir4, Net1	SIRT1-7	PRC4
	Hst1	Set3, Sum1		
	Hst2	Desconegut		

Quina és la relació entre els nivells d'acetilació i l'activitat transcripcional?

Aquesta pregunta no troba una resposta única en la literatura existent. Una primera resposta assumeix que són els nivells globals d'acetilació en lisines els que permeten un estat transcripcionalment actiu o activable. Aquesta hipòtesi es sustenta en el fet que els residus de lisina no metilats tenen càrrega positiva, fet que els permet una interacció més forta amb el DNA, el qual està carregat negativament. L'addició d'un grup metil resta carrega neta per cada lisina, de manera que la unió entre el DNA i la cua de la histona es relaxa (Vettese-Dadey et al. 1996). Aquesta hipòtesi ha estat comprovada en llevats, arribant a la conclusió que és certa per a determinats residus de la histona H4. Així doncs, experiments en els què es substitueix la histona H4 endògena per una histona H4 amb mutacions en els residus de lisina 5, 8 i 12 que mimetitzen un estat acetilat, demostren que l'efecte de la metilació d'aquests residus té un efecte acumulatiu, no específic, sobre l'expressió d'uns 1200 gens. Com major és el nombre de lisines acetilades, majors són els nivells d'expressió. En canvi, l'acetilació de la lisina 16 en la mateixa histona H4 és representativa d'un segon model de regulació. En aquest cas, la presència de lisina 16 acetilada correlaciona amb l'estat transcripcionalment actiu d'un centenar de gens, independentment de l'estat d'acetilació dels altres residus (Dion et al. 2005).

Aquests resultats mostren dos models diferents de com l'acetilació altera l'expressió dels gens: el primer model diu que l'acetilació de les histones té un efecte global de relaxació de la cromatina per supressió de les càrregues positives en les histones mentre que el segon model presenta l'acetilació de determinats residus com a marques específiques que poden ser reconegudes per activitats que alteren l'arquitectura de la cromatina. L'existència de dominis proteics específics de reconeixement de lisines acetilades, els bromodominis, dona suport a aquest segon model i a l'existència d'un veritable codi d'histones. Aquests dominis de 110 aminoàcids es troben en múltiples proteïnes associades a la cromatina (Taula 5) i van ser descoberts per primer cop en el coactivador transcripcional Brahma de *Drosophila melanogaster* (Tamkun et al. 1992), tot i que la seva estructura va ser resolta anys més tard en el factor PCAF (Dhalluin et al. 1999).

Taula 5. Proteïnes que contenen bromodominis i la seva funció.

Família	Exemples	Organisme
HAT	Gcn5 p300/CBP PCAF (TAF) _{II} 250	Llevats > humans Varis multicel·lulars Humans, ratolí Llevats > humans
Remodeladors cromatina	Brahma Swi2 Snf2 Brg1	<i>D.melanogaster</i> Humans
BET	Bdf1 Bdf2 Bdr4 Brd2	Humans

Entre les proteïnes que tenen bromodominis hi ha proteïnes HAT, complexes remodeladors de cromatina depenents d'ATP i la més desconeguda família de reguladors transcripcionals BET (Bromodomini i ET domini) (Taula 5). En tots els casos, les proteïnes que reconeixen lisines acetilades participen en l'activació a través de dos mecanismes: l'addició de noves marques de acetilació (en el cas de HATs) o bé la remodelació de la cromatina per crear entorns de cromatina més accessible i activa.

L'activació del gen interferó beta (IFN- β) després de la infecció vírica és un clar exemple de com un patró específic de PTMs és reconegut de forma seqüencial per complexes remodeladors de cromatina i maquinària transcripcional (Agalioti et al. 2000)(Figura 12). El gen IFN- β té un *enhancer* flanquejat per dos nucleosomes, un dels quals tapa la caixa TATA i el lloc d'inici de transcripció o TSS. La seqüència d'esdeveniments que condueixen a l'activació transcripcional comença pel reconeixement, a nivell genètic, de l'*enhancer* del gen IFN- β per part d'un dels tres factors de transcripció que hi tenen seqüències de reconeixement (NF- κ B, membres de la família IRF i l'heterodímer ATF-2/c-Jun)(Figura 12a). Aquest fet permet el reclutament del complex amb activitat HAT GCN5/PCAF, el qual acetila el nucleosoma que hi ha sobre el promotor i que impedeix la formació del complex de preiniciació o PIC (Figura 12b). Seguidament entra sobre el promotor el complex de CBP/RNA polimerasa II, el qual es seguit pel reclutament, via interaccions amb CBP i les lisines acetilades (a través del seu bromodomini) del complex remodelador de cromatina SWI/SNF (Figura 12c). Aquest fet, altera l'estructura del nucleosoma per mitjà de mecanismes poc coneguts, el que permet el reconeixement de la caixa TATA per part de la subunitat TBP del factor de transcripció TFIID. El reconeixement de la caixa TATA per part de TBP provoca que el DNA es doblegui de tal manera que el nucleosoma que hi havia sobre la caixa TATA i el TSS llisqui a una nova localització 36 parell de bases riu avall (Figura 12d), fet que permet l'inici de la transcripció.

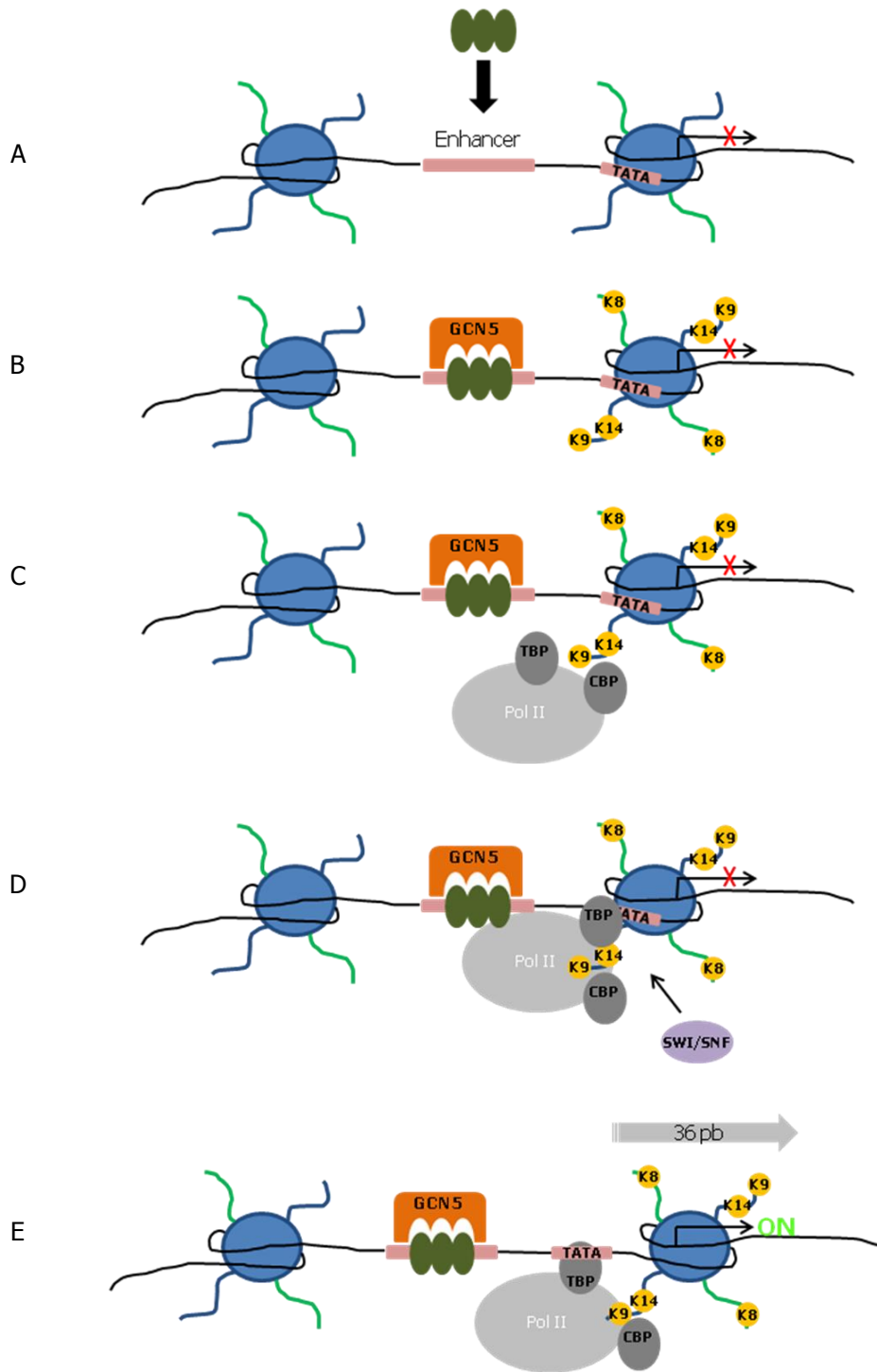


Figura 12. Seqüència d'esdeveniments genètics i epigenètics en l'activació del gen IFN- β .

L'activació del gen IFN- β s'inicia quan algun dels seus activadors (esferes verdes) reconeix el seu *enhancer* (A), el qual es troba situat entre dos nucleosomes (cercles blaus). Aquest fet provoca el reclutament del complex HAT GCN5, el qual acetila les posicions de lisina 9 i 14 en la cua de la histona H3 (línies blaves) i la posició lisina 8 en la cua de la histona H4 (línies verdes), entre d'altres (B). Les lisines acetilades (cercles grocs) són reconegudes pel bromodomini de CBP, fet que permet el reclutament del complex de preiniciació amb la RNA pol II (Pol II)(C). La presència del complex de preiniciació facilita el reclutament del complex remodelador de cromatina SWI/SNF el qual modifica el nucleosoma de tal manera que la subunitat TBP de reconeixement de la caixa TATA s'hi pot unir (D). La unió de TBP a la caixa TATA provoca el lliscament del nucleosoma 36 pb riu avall, fet que comporta l'inici de la transcripció (E).

Treballs posteriors duts a terme pel mateix grup demostren que determinats residus de lisina acetilats són essencials per tal reclutament seqüencial de determinades activitats. Així, l'acetilació de la lisina 8 és necessària per al reclutament dels complexos remodeladors de cromatina SWI/SNF, mentre que les lisines acetilades 9 i 14 de la histona H3 són essencials per al reclutament de TFIID i per tant, del reconeixement de la caixa TATA (Agalioti et al. 2002).

El fet que nivells d'acetilació elevats en determinades lisines són reconeguts per complexos remodeladors de cromatina posa de manifest una associació que havia estat feta molts anys enrere en el locus de la β -globina de cèl·lules eritroides de pollastre per mitjà de mètodes indirectes. En aquests estudis, Weintraub i col·laboradors varen veure per primer cop que la cromatina de gens actius és més accessible a la fragmentació per diferents endonucleases que la cromatina de gens inactius, demostrant una relació entre l'activitat transcripcional i l'accessibilitat de la cromatina (Weintraub and Groudine 1976). Tot plegat posa de manifest la relació positiva que existeix entre els nivells d'acetilació d'un determinat gen, una conformació de la cromatina oberta i accessible, i el seu estat transcripcional actiu.

La metilació de les histones també configura un codi

La metilació de les cues de les histones es pot donar en diferents graus. Així podem trobar lisines i arginines (parlarem bàsicament de lisines) mono, di o trimetilades. A més a més la metilació de les histones s'associa, depenent del residu implicat, amb activació o bé inactivació transcripcional (Taula 2)(Revisat en (Kouzarides 2007; Kouzarides and Berger 2007; Martin and Zhang 2005)). Aquestes dues característiques afegeixen un extra de complexitat a l'estudi de la metilació de les histones, complexitat que es veu en part reduïda pel gran nombre de treballs que comencen a mostrar els mecanismes moleculars a la base de les funcions d'aquesta PTM.

Com en el cas de l'acetilació, la metilació de lisines en diferents posicions i diferents graus conforma un codi que pot ser interpretat per proteïnes capaces de llegir la senyal i coordinar el reclutament de factors amb diferents activitats. En el cas de les lisines metilades, s'han trobat tres dominis capaços de reconèixer específicament aquests residus metilats: són el cromodomini, el domini tudor i la repetició WD40, els quals trobem en diferents proteïnes que reconeixen diferents residus de lisina metilats (Figura 13) (Revisat en (Kouzarides and Berger 2007; Martin and Zhang 2005)). Les conseqüències de la unió d'aquestes proteïnes són tant a nivell d'expressió com d'estructura de la cromatina i reparació del DNA.

Encara que com hem vist, existeixen múltiples residus metilables, ens centrarem bàsicament en tres: una posició associada positivament amb la transcripció com és la lisina 4 i dues posicions associades a silenciament com són la lisina 9 i la lisina 27. En cada cas veurem les activitats enzimàtiques que metilen aquestes lisines (resumides en la taula 6) i els mecanismes moleculars de reconeixement de la posició metilada i les seves conseqüències.

d'heterocromatina depèn de la producció de RNAs petits d'interferència (siRNA)(Hall et al. 2002; Volpe et al. 2002). Segons aquest model, la producció de siRNA provinents de regions heterocromàtiques acaben amb el reclutament de SUV39H1 i HP1 (Per a una extensa revisió dels mecanismes moleculars que condueixen a la formació d'heterocromatina veure (Grewal and Jia 2007)).

Més evidències del paper de la metilació de la lisina 9 en el silenciament provenen de l'observació que gens silenciats en regions eucromàtiques presenten un enriquiment per lisina 9 metilada en els seus promotors, sobre tot les formes mono i dimetilades (la forma trimetilada queda més aviat restringida a regions heterocromàtiques)((Peters et al. 2003; Rice et al. 2003); Revisat en (Kouzarides and Berger 2007; Martin and Zhang 2005)). En la mateixa línia, ratolins que són deficientes per les HKMTs de lisina 9, G9a o G9a-related protein (GLP), tenen molt reduïts els nivells de metilació de la lisina 9 en regions eucromàtiques (Tachibana et al. 2005). En alguns casos coneixem quin és el mecanisme que direcciona aquestes HKMTs a regions eucromàtiques, com per exemple la proteïna retinoblastoma (Rb), la qual silencia gens de la fase S reclutant SUV39h1 i SUV39h2, fet que culmina amb el reclutament de HP1 sobre els mateixos promotors (Nielsen et al. 2001).

Com veurem en l'apartat d'epigenètica i càncer, una bona part dels coneixements que tenim sobre el rol la lisina 9 metilada en la regulació epigenètica provenen del seu paper de silenciament associat a la metilació del DNA en càncer.

Memòria cel·lular i la metilació de les lisines 4 i 27.

La memòria cel·lular és el fenomen per mitjà del qual una cèl·lula és capaç de recordar un determinat estat transcripcional quan els estímuls que han provocat aquest estat transcripcional ja no hi són. El més ben conegut sistema de memòria el conformen dos sistemes antagonistes, el grup de gens trithorax (Trx) i el grup de gens polycomb (PcG), la funció més important dels quals és la regulació dels patrons d'expressió dels gens homeobox (Hox). Aquests gens, essencials per al desenvolupament, confereixen la identitat que pertoca a cada un dels segments al llarg de l'eix anteroposterior en metazous (Pearson et al. 2005). Aquest sistema no només permet l'establiment, sinó també el record a través de la divisió cel·lular, dels patrons d'expressió dels gens Hox que conferiran la identitat dels segments, fet que conforma un veritable sistema de memòria cel·lular.

En la dècada dels anys 40 varen ser descoberts dos mutants en *Drosophila* amb un fenotip característic. Aquests mutants es caracteritzaven per l'aparició de pintes sexuals en totes les potes del mascle, enlloc de sortir només en el primer parell de potes on realment pertanyen. Aquests mutants es varen anomenar, en al·lusió al fenotip que provocaven, extra sex combs (Esc) i polycomb (Pc), gen aquest últim, que ha acabat donant nom a tota la família PcG. Aquest fenotip es causat per la transformació de la identitat d'un segment (en aquest cas varis segments) del cos de *Drosophila* en la identitat d'un altre, com a conseqüència de la desrepressió d'alguns gens Hox en aquests segments per les mutacions en els gens Pc i Esc. Els

gens Trx varen ser descoberts cap a la dècada dels 80 com a supressors del fenotip imposat per la mutació Esc (Revisat en (Grossniklaus and Paro 2007; Ringrose and Paro 2004)).

Anàlisis moleculars recents han permès conèixer que ambdós grups realitzen la seva funció modificant la cromatina per a mantenir el silenciament (PcG) o l'activitat (Trx) dels seus gens diana. Aquesta activitat modificadora de la cromatina la duen a terme a través de complexos multiproteics amb activitat HKMT de la lisina 27 (PcG) i de la lisina 4 (Trx) en la histona H3 (Revisat en (Kouzarides and Berger 2007; Martin and Zhang 2005; Ringrose and Paro 2004)). Vegem doncs aquestes dues PTMs i la seva funció en la regulació transcripcional.

Metilació de la lisina 4

L'addició d'aquesta marca pot ser catalitzada per diferents HKMTs (resumides en la taula 7), com Set1 en llevats o les proteïnes mixed-lineage leukemia (MLL) en humans, ambdues homòlogues de Trithorax (Trx) de *Drosophila*. La metilació de la lisina 4 ha estat llargament associada a activació transcripcional i la trobem en regions eucromàtiques, sobre gens actius o activables. Els diferents estats de metilació (mono, di o trimetilació) tenen distribucions diferents i han estat associats a diferents estats de transcripció. Estudis a nivell global en llevat demostren que la trimetilació la trobem enriquida en la regió promotora, la dimetilació es troba bàsicament en el cos del gen i la monometilació la trobem cap al final del gen (Bernstein et al. 2002; Pokholok et al. 2005)(Revisat en (Kouzarides and Berger 2007; Martin and Zhang 2005)). Resultats similars han estat obtinguts tant en *Drosophila* (Schubeler et al. 2004), com en ratolí i humans (Bernstein et al. 2005).

Encara que poques, es comencen a conèixer algunes de les proteïnes que reconeixen aquesta marca i quins efectes tenen per a l'activació transcripcional, com són les proteïnes CHD1 i WDR5 (Figura 13). CHD1 s'ha descrit com a component del complex HAT anomenat SAGA (Spt-Ada-Gcn5-acetyltransferase) el qual regula l'activitat d'un gran nombre de gens en llevat. Així doncs, el complex activador SAGA pot ser reclutat, a través del seu component CHD1 sobre lisines metilades (Revisat en (Kouzarides and Berger 2007; Martin and Zhang 2005)). La segona proteïna identificada per unir lisina 4 metilada, WDR5, forma part del complex H3K4 metiltransferasa MLL (Mixed-lineage leukemia). De forma interessant, WDR5 s'uneix preferentment a la lisina 4 dimetilada fet que suggereix que aquesta proteïna podria guiar a aquest complex HKMT a regions amb K4 dimetilades per a donar lloc a regions trimetilades (Revisat en (Martin and Zhang 2005)).

Metilació de la lisina 27

Aquesta marca de metilació associada a silenciament transcripcional es dipositada sobre les histones per membres de PcG. Com ja hem dit, les principals dianes d'aquest silenciament mediat per PcG són els gens Hox, tot i que el silenciament mediat per la metilació de la lisina 27 ha estat relacionat també amb la inactivació del cromosoma X, l'imprinting i el desenvolupament del càncer (Revisat en (Cao and Zhang 2004)).

La proteïna que catalitza la metilació d'aquest residu en mamífers és EZH2, homòloga de la proteïna de *D.melanogaster* enhancer of zeste (E(Z)). Bioquímicament s'han caracteritzat dos classes de complexos PcG: els polycomb repressor complex 1 i 2 (PRC1 i PRC2). Tot i que aquests són els dos principals complexos, estudis recents n'han identificat d'altres, que com veurem més endavant, juguen un paper important en processos patològics com el càncer.

El polycomb repressor complex 2 (PRC2)

Diferents complexos PRC2 que contenen els productes dels gens Esc-E(z) (o els homòlegs de mamífers Eed-Ezh2) han estat aïllats a partir de *Drosophila* (Czermin et al. 2002; Muller et al. 2002) i cèl·lules de mamífer (Cao et al. 2002; Kuzmichev et al. 2002). Aquests complexos comparteixen tots els productes de 4 gens principals: enhancer of zeste (E(z)), Extra sex combs (Esc), suppressor of zeste 12 (Su(z)12) i p55. En la taula 6 veiem resumits aquests components així com els homòlegs d'humans i les seves funcions, si es coneixen.

Un dels components clau del PRC2 és la seva activitat HKMT, la qual es deu a EZH2. Aquesta proteïna no té activitat sinó és en el context del complex PRC2 (Czermin et al. 2002; Muller et al. 2002). Tot i que només ens centrarem en la capacitat de EZH2 de metilar la lisina 27, aquesta HKMT pot metilar, dins el context del PRC2, la posició de lisina 26 en la histona H1 (Revisat en (Schwartz and Pirrotta 2007)).

El polycomb repressor complex 1 (PRC1)

El complex PRC1 va ser aïllat en *Drosophila* i es compon dels productes de 4 gens principals: Polycomb (Pc), Posterior sex combs (Psc), Polyhomeotic (Ph) i Ring. El complex homòleg de mamífers va ser aïllat a partir de cèl·lules HeLa i és molt similar al complex prèviament caracteritzat de *Drosophila* (Revisat en (Schwartz and Pirrotta 2007)). La taula 6 mostra els components d'aquest complex així com els seus homòlegs humans i la seva funció, si és coneguda.

Podríem dir que la proteïna principal d'aquest PRC1 és Pc, la qual presenta un domini del tipus cromodomini que és capaç de reconèixer lisina 27 trimetilada. Segons aquest model, PRC2 és el complex que marca les histones sobre les quals serà reclutat el PRC1 a través del seu cromodomini. Els mecanismes que condueixen al silenciament són encara poc coneguts tot i que requereixen tant dels membres del PRC2 com del PRC1 (Revisat en (Schwartz and Pirrotta 2007)).

Altres PRCs: PRC3 i PRC4

Estudis més recents han permès identificar altres complexos PRC, els quals tenen una composició molt similar al PRC2. El complex PRC3 es caracteritza per contenir isoformes diferents del gen Eed (per inici alternatiu de la traducció a partir del mateix mRNA) anomenades Eed2 i Eed3 (Kuzmichev et al. 2004). Els mateixos investigadors descriuen en un treball posterior l'existència d'un PRC4, el qual a més de contenir la isoforma Eed2 conté també

la HDAC NAD depenent SirT1. El PRC4 pot metilar la lisina27, tot i que en presència de la histona H1 en mostra preferència per la lisina 26 (Kuzmichev et al. 2005).

Taula 6. Principals components dels complexos polycomb PRC1 i PRC2

Complex	Components <i>Drosophila</i>	Homòlegs Humans	Funció
PRC1	Pc	Cbx2/4/8	Unió de PRC1 a K27 metilada
	Ph	Edr1/2/3	Desconeguda
	Psc	Bmi1	Incrementa l'activitat de RING
	Ring	hRing	E3 ubiquitin ligasa de K119 en histona H2A
PRC2	E(z)	Ezh2	Metilació de la lisina 27
	Esc	Eed	Cofactor de EZH2
	Su(z)12	SUZ12	Desconeguda
	p55	RbAp46/48	Desconeguda

Complexitat dels patrons de modificacions de histones: els dominis bivalents

La gran capacitat de què disposem avui en dia per a realitzar estudis globals a escala genòmica ha permès obtenir mapes globals de modificacions d'histones i de la distribució d'altres proteïnes al llarg del genoma de diferents organismes model com el llevat, el ratolí i l'home. Sorprenentment, l'estudi de la distribució d'histona H3 trimetilada en lisina 4 i 27 en ESCs murines ha demostrat l'existència de dominis cromatínics en els què les dues marques coexisteixen, en el què els autors denominen dominis bivalents (Bernstein et al. 2006). Aquests dominis s'estenen sobre les regions promotores de gens clau regulats durant el desenvolupament els quals globalment presenten nivells baixos d'expressió. Al llarg de la diferenciació, els dominis bivalents tendeixen a resoldre's cap a promotors enriquits per lisina 4 trimetilada, en el cas de gens actius, o bé enriquits per lisina 27 trimetilada, en el cas de gens silenciats. Aquests experiments mostren com la totipotencialitat de les ESC ve determinada per patrons de cromatina específics que tant permeten a la cèl·lula activar o inactivar determinats gens durant la diferenciació.

Modulació dels nivells de metilació: Les metiltransferases i desmetilases d'histones

Ja hem vist algunes de les HKMTs que catalitzen la metilació de residus concrets com la K4, K9 i K27 tot i que aquestes no són però les úniques posicions de lisines metilables, ni les lisines són els únics residus metilables (Figura 11 i taula 2). La taula 7 conté un resum de les principals proteïnes amb activitat HKMT (Revisat en (Kouzarides 2007; Kouzarides and Berger 2007; Martin and Zhang 2005)). El domini proteic que els confereix l'activitat metiltransferasa està conservat evolutivament i s'anomena SET, nom que prové de les inicials de les primeres proteïnes amb activitat HMT descrites: SU(VAR)3-9, Enhancer of zeste i TriThorax (Tschiersch et al. 1994).

La marca de metilació dipositada per aquestes HKMT és relativament estable, fet que va deixar en la incertesa l'existència de proteïnes capaces de esborrar aquesta marca. El descobriment l'any 2004 que la proteïna LSD1, prèviament identificada en complexos amb activitat HDAC (Revisat en (Martin and Zhang 2005)) és una desmetilasa de la lisina 4 en la histona H3, va acabar amb el dubte (Shi et al. 2004). Posteriorment s'han trobat altres proteïnes, amb un domini catalític diferent al de LSD1 anomenat jumonji C (JmjC), que són capaces de desmetilar posicions com la lisina 9 i la 36. En aquest cas, el fet que la metilació de les histones estigui associada tant a activació com a silenciament transcripcional (Taula 2 i Figura 13) fa que la desmetilació també hi estigui associada, depenent dels residus afectats. Aquest fet contrasta amb la desacetilació de les histones, la qual no s'ha vist associada a activació transcripcional en cap cas. Molt recentment s'han trobat desmetilases de la lisina 27 que contenen dominis JmjC. La desmetilasa UTX participa de la regulació de gens HOX i s'associa amb complexos tritòrax (Lee et al. 2007). La desmetilasa de lisina 27 Jmjd3, relacionada amb UTX, està també associada a activació transcripcional i a la regulació de l'eix anteroposterior en animals (De Santa et al. 2007; Lan et al. 2007; Xiang et al. 2007). La taula 7 conté les proteïnes amb activitat desmetilasa conegudes així com els residus que poden desmetilar.

Taula 7. Classificació de les diferents activitats metiltransferasa i desmetilasa conegudes

Posició	Metiltransferasa	Organisme	Posició	Desmetilasa	Organisme
H3K4	MLL1-5	Mamífers	H3K9	JHDM2a/2b	Mamífers
	SET1A/1B			JMJD2A/2B/2C/2D	
	ASH1			JHDM3A	
	SET1	Llevat		GASC1	
H3K9	SUV39H1/H2	Mamífers	H3K4	LSD1/BHC110	Mamífers
	G9a		H3K36	JHDM1a/1b	Mamífers
	ESET/SETDB1			JMJD2A	
	EuHMTase/GLP			JHDM3A	
	CLL8			JMJD2C	
	RIZ1			GASC1	
Clr4	Llevat	H3K27	UTX	Mamífers	
H3K27	EZH2	Mamífers	Jmjd3		
H3K36	SET2	Mamífers	<hr/>		
	NSD1				
	SYMD2				
	SET2	Llevat			
H3K79	DOT1	Mamífers			
	DOT1	Llevat			
H4K20	Pr-SET 7/8	Mamífers			
	SUV4 20H1/H2				
	Set9	Llevat			

Un model conjunt de regulació epigenètica: la metilació del DNA i les modificacions de les histones

Fins ara hem vist per separat les modificacions epigenètiques que es poden donar sobre els dos components de la cromatina, la metilació del DNA i les modificacions de les histones, però existeix una qüestió fonamental que fins ara no hem abordat: quina és la relació entre ambdues modificacions epigenètiques? En aquest sentit existeixen múltiples estudis que demostren que ambdós nivells de regulació de l'activitat transcripcional estan íntimament relacionats, tot i que amb resultats moltes vegades contradictoris. Bàsicament podem plantejar dos models no excloents:

Un primer model és el que planteja que la metilació del DNA és l'esdeveniment que inicia el silenciament transcripcional. Les evidències que donen suport a aquesta hipòtesi provenen bàsicament d'experiments fets amb transgens. Treballs realitzats per Eden i col·laboradors demostren que el DNA integrat del transgèn timidilat kinasa (tk) s'associa a histones hipoacetilades quan el transgèn ha estat prèviament metilat *in vitro*, mentre que el transgèn no metilat s'associa amb histones hiperacetilades (Eden et al. 1998). En un experiment similar, Schubeler i col·laboradors, usant una construcció formada pel promotor de la β -globina i el gen reporter de proteïna fluorescent verda (GFP), demostren que la integració de la còpia del transgèn metilada *in vitro* no només s'associa amb silenciament transcripcional i histones hipoacetilades sinó que a més a més la cromatina és més inaccessible a la digestió amb endonucleases. Per contra, la integració de la còpia no metilada es manté activa en un entorn accessible i hiperacetilat (Schubeler et al. 2000). Treballant amb la línia de CCR HCT116, Espada i col·laboradors han demostrat que els nivells de dimetilació i trimetilació de la lisina 9 en la histona H3 depenen de la metilació del DNA. L'eliminació de la funció de DNMT1 en aquest model de càncer provoca la hipometilació de regions de DNA repetitiu i un decrement global dels nivells de metilació de l'esmentat residu de lisina (Espada et al. 2004).

Segons aquests resultats, la metilació del DNA és capaç de provocar el silenciament transcripcional a través de la modificació de l'estructura de la cromatina, possiblement reclutant proteïnes amb activitat HDAC. En suport d'aquesta hipòtesi, s'ha vist que proteïnes amb dominis de unió a 5mC, les anomenades MBDs (veure apartat de Metilació del DNA) poden interaccionar directament amb proteïnes amb activitat HDAC. És el cas de MeCP2, la qual s'ha vist que pot formar complexos, a través del seu domini TRD, amb el complex repressor amb activitat HDAC anomenat Sin3 (Taula 4), tant en cèl·lules de mamífer (Nan et al. 1998) com en cèl·lules de *Xenopus* (Jones et al. 1998). Altres MBDs, com MBD2 i MBD3 estan involucrades en el reclutament del complex amb activitat HDAC Mi-2/NuRD (Taula 4) a regions de DNA metilat (Wade et al. 1999).

Un segon model és el que assumeix que les modificacions de la cromatina són les que inicien el silenciament. Aquest model implica que la metilació del DNA és un esdeveniment secundari al silenciament transcripcional, precedit per canvis a nivell de les histones. Així doncs la metilació del DNA afectaria a gens prèviament silenciats o amb una expressió molt reduïda (Revisat en (Bird 2002)). Les evidències més fortes provenen de dos organismes model una mica allunyats

de mamífers. En el fong *Neurospora crassa*, mutacions en la metiltransferasa que catalitza la metilació de la lisina 9 en la histona H3 provoquen la pèrdua de la metilació genòmica (Tamaru and Selker 2001). Treballs posteriors demostren que la metilació del DNA és totalment dependent dels nivells de trimetilació, i no de dimetilació, de la lisina 9 en el mateix fong (Tamaru et al. 2003). De forma similar, els nivells de metilació del DNA en la planta *Arabidopsis thaliana* depenen de la metiltransferasa de la histona H3 KRYPTONITE (Jackson et al. 2002) i de l'activitat de DDM1, membre de la família de remodeladors de cromatina Swi2/Snf2 (Vongs et al. 1993). Mutacions en Lsh1, l'homòleg en mamífers de DDM1, causen hipometilació en ratolí de forma similar a la hipometilació induïda en mutants DDM1 en *A.thaliana* (Geiman et al. 2001). Altres estudis amb el transgèn tk han suggerit que el silenciament induït per la metilació depèn del reclutament de proteïnes de la cromatina sobre la regió a silenciar (Kass et al. 1997).

Dues evidències ens indiquen que aquest segon model podria ser el més rellevant en mamífers. En primer lloc, en aquests organismes la metilació del DNA sol ser un esdeveniment que afecta a gens amb baixa expressió o prèviament silenciat (per exemple, gens metilats i silenciat en càncer). En segon lloc, els nivells més alts de metilació del DNA es donen en regions de DNA pericentromèric, altament enriquides per histona H3 amb lisina 9 trimetilada (Revisat en (Bird 2002)).

L'existència d'aquests dos models en cap cas podem afirmar que sigui exclouent. Més aviat hauríem de dir que es pot establir un diàleg multidireccional entre els diferents factors que afecten a l'estructura de la cromatina i l'expressió gènica (Figura 14).

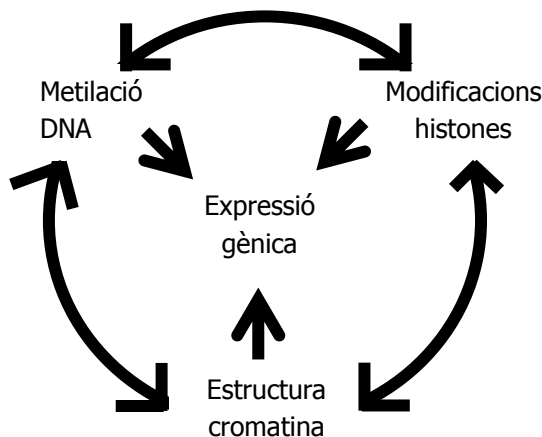


Figura 14. Diàleg entre els diferents components de la regulació epigenètica. Més que una relació lineal entre les diferents capes de regulació epigenètica de l'activitat gènica, s'estableix una relació circular en la que cada un dels components té efectes sobre els altres dos i sobre la transcripció, la qual està en el centre d'aquesta regulació.

4. Alteracions epigenètiques en càncer

Fins ara hem vist alguns dels mecanismes epigenètics que operen en una cèl·lula normal i quins són els mecanismes moleculars a través dels quals duen a terme la seva funció reguladora. En ocasions però, aquests sistemes no funcionen o ho fan de forma anormal, donant lloc a l'aparició de malalties com el càncer.

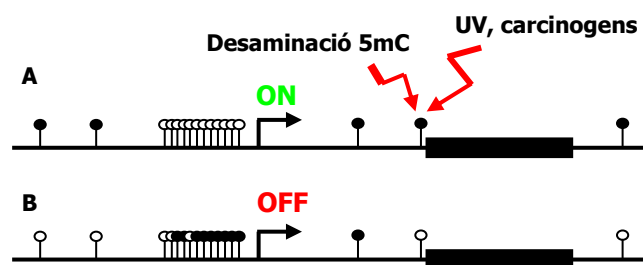
Al llarg de l'apartat d'epigenètica hem vist en primer lloc la metilació del DNA i en segon lloc les modificacions de les histones. Finalment hem pogut integrar en un sol model els dos mecanismes de regulació. L'ordre que hem seguit en la introducció d'aquests mecanismes no és fortuït, sinó que segueix un cert paral·lelisme a la recerca epigenètica realitzada en càncer durant els passats més de 25 anys.

Aquesta recerca, iniciada a principis dels anys 80, es va centrar des del primer moment en les alteracions de la metilació en tumors, tant a nivell de contingut global com a nivell dels patrons de metilació. Ben aviat es va veure que les pertorbacions de la metilació del DNA en tumors tenen dues conseqüències clarament diferents i oposades: la hipometilació o pèrdua de metilació global del genoma i la metilació de illes CpG, les quals estan normalment lliures de metilació en teixits sans (Revisat en (Baylin 2005; Ehrlich 2002; Esteller 2007; Feinberg and Tycko 2004; Herman and Baylin 2003; Jones and Baylin 2002; Jones and Baylin 2007)). Aquest interès per les alteracions de la metilació del DNA en tumors va centrar la major part de la recerca fins a finals dels anys 90, quan la cerca de PTMs de les histones associades a gens hipermetilats i silenciats en tumors va entrar en escena amb una major força. Finalment el cercle es comença a tancar. La recerca dels dos darrers anys comença a deixar entreveure una hipòtesi molt atractiva segons la qual els patrons aberrants de metilació del DNA en càncer depenen bàsicament dels patrons de modificacions d'histones. Tot i que com ja hem vist els dos sistemes de regulació estan íntimament relacionats, encara està per veure quin grau de dependència hi ha entre ells.

En aquest apartat final de la introducció veurem com la desregulació epigenètica contribueix a través de diferents mecanismes al procés tumorigènic: la mutagènesi induïda per la 5mC, la hipometilació i la hipermetilació (Figura 15).

Figura 15. Principals alteracions epigenètiques en càncer.

L'estat desmetilat de dinucleòtids CpG (cercles) en la regió promotora d'un gen (punta de fletxa sobre línia horitzontal) en determinen el seu estat transcripcional actiu (ON). Els dinucleòtids CpG en la regió codificant (els exons apareixen com a caixes negres) estan metilats (cercles negres), fet que els fa més susceptibles a patir mutacions per



desaminació espontània o induïda per llum ultraviolada o carcinògens externs (A). D'altra banda, la presència de metilació en la illa CpG (B) provoca el silenciament del gen (OFF). A més a més, nucleòtids que estan normalment metilats fora de la illa CpG, ara els trobem desmetilats, fet que pot comprometre l'estabilitat del genoma.

Mutagènesi associada a la metilació del DNA

Ja hem vist com la sola presència de 5mC pot provocar diverses alteracions a nivell de la seqüència del DNA, fenomen responsable, entre d'altres, de la desaparició de dinucleòtids CpG en genomes metilats. Si les mutacions es donen en regions crítiques per al funcionament dels

gens, com regions codificants o promotors, poden acabar per malmetre o eliminar la funció del gen (Figura 15). Algunes estimacions indiquen que aquest mecanisme és el responsable de fins al 50% de les mutacions puntuals en el gen supressor tumoral p53, i pot contribuir fins a un 30% del total de mutacions causants de malaltia (Rideout et al. 1990). De forma similar, mutacions en p53 es veuen incrementades en posicions de 5mC tant per la llum ultraviolada (UV) com per carcinògens externs, als quals la 5mC és més sensible (Revisat en (Jones and Baylin 2002)).

La hipometilació en tumors

La hipometilació que mostren alguns tumors va ser la primera alteració epigenètica descrita en càncer. Tot i aquests descobriments inicials, la recerca encaminada a entendre quin podia ser el seu paper en el procés tumorigènic va caure una mica en l'oblit, produint pocs resultats. A aquesta situació hi va ajudar el fet que la majoria de tècniques de rastreig de la metilació que s'anaven desenvolupant tenien un clar biaix cap a l'aïllament de regions metilades *de novo* (bàsicament illes CpG), les quals tenien un paper més clar en el silenciament de gens supressors tumorals. Malgrat això, avui podem entendre una mica més quina és la relació d'aquest fenomen amb el càncer.

La hipometilació és una comparació relativa entre un estat inicial metilat i un estat final parcial o totalment desmetilat. Per entendre les diferents situacions en les què la hipometilació pot jugar un paper en càncer primer hem de recordar quines regions del genoma es poden desmetilar, és a dir, quines regions estan metilades de forma normal. Aquestes són bàsicament dues: una part majoritària del genoma formada per seqüències i elements repetitius i una porció minoritària de illes CpG associades a gens silenciats del cromosoma X inactiu, gens imprintats i gens metilats específics de teixit. Per tant, en càncer la hipometilació pot ser de dues menes:

1. global: aquella que afecta bàsicament a elements repetitius al llarg de tot el genoma. Aquesta desmetilació s'ha associat principalment a la inestabilitat genètica.
2. regional: aquella que afecta regions discretes, bàsicament illes CpG metilades. Aquesta desmetilació s'ha associat a la reactivació transcripcional de gens silenciats durant el desenvolupament i diferenciació.

A continuació veurem aquests dos fenòmens i quines conseqüències tenen.

Hipometilació regional

L'any 1983 Andrew Feinberg i Bert Vogelstein, per mitjà de l'anàlisi de *southern blot* de DNA tallat amb enzims de restricció sensible a la metilació varen descriure per primera vegada la hipometilació de gens específics en tumors primaris humans, entre ells el gen Hras (Feinberg and Vogelstein 1983a; Feinberg and Vogelstein 1983b). Avui coneixem varis gens en diferents tipus de càncer que desmetilen durant el procés tumoral per a produir un guany de funció.

Entre ells hi ha gens imprimats i gens específics de teixit. El primer cas conegut de pèrdua d'imprinting o *loss of imprinting* (LOI) en tumors és el que afecta al locus IGF2/H2 en tumors de Wilms (tumors sòlids més freqüents en la infantesa), fet que provoca l'expressió bial·lèlica del factor de creixement autocrí IGF2. Posteriorment, LOI d'aquest locus s'han trobat en altres tipus de càncer com el de pulmó, mama, ovari i en gliomes (Revisat en (Feinberg 2007; Feinberg and Tycko 2004)).

També es coneix gens metilats, no imprimats, que desmetilen el tumors, com el cas que ja hem vist de Hras, conjuntament amb la ciclina D2 i el gen Maspín en càncer gàstric; la anhidrasa carbònica IX en càncer renal, el gen S100A4 en CCR i el gen HPV16 en càncer cervical (Revisat en (Ehrlich 2002; Feinberg 2007; Feinberg and Tycko 2004)). És també el cas dels gens de l'anomenat grup MAGEA1, coneguts com a antígens de càncer i testicles. La seva desmetilació en tumors i línies cel·lulars correlaciona amb la seva reexpressió (De Smet et al. 1996).

Hipometilació global

Aquest fenomen va ser observat per primera vegada en carcinomes hepatocel·lulars químicament induïts, els quals mostraven nivells globals de metilació inferiors als del teixit normal (Lapeyre and Becker 1979). Més tard, treballs duts a terme en el laboratori de Melanie Ehrlich (una de les investigadores pioneres en aquest camp) varen estimar els nivells globals de 5mC en tumors primaris humans per mitjà de HPLC (*High-performance liquid chromatography*). Els resultats obtinguts indicaven que en la progressió de tumor benigne a maligne i metastasi, els nivells globals de 5mC disminuïen progressivament, en comparació amb diferents teixits somàtics postnatsals (Gama-Sosa et al. 1983a). Una mica més tard, Andrew Feinberg, en col·laboració amb la mateixa Ehrlich i Charles Gehrke varen poder fer la primera comparació entre els nivells de metilació de tumors colorectals i la mucosa sana a partir de la qual s'originen, obtenint com a resultat uns nivells de hipometilació dels tumors entre el 8 i el 10% (Feinberg et al. 1988). Avui sabem que la hipometilació genòmica global afecta a molts tipus de tumors diferents, entre els quals hi trobem tumors de pròstata metastàsics en comparació amb la pròstata normal; leucòcits de leucèmia limfocítica crònica de cèl·lules B en comparació amb leucòcits normals; carcinomes hepatocel·lulars en comparació amb teixit de fetge normal i en carcinomes ductals de mama en comparació amb l'epiteli normal (Revisat en (Ehrlich 2002)).

La seqüenciació quasi definitiva del genoma humà ens ha permès conèixer que aquest està constituït en un 45% per elements repetitius (Lander et al. 2001). Aquest component repetitiu és el que acumula uns nivells de 5mC més elevats en el genoma humà, fet abastament demostrat per múltiples treballs en la majoria de teixits somàtics ((Rollins et al. 2006); Revisat en (Yoder et al. 1997)). Gama-Sosa i col·laboradors varen ser els primers en estudiar el contingut de 5mC en seqüències repetides Alu, així com la família de DNA altament repetitiu que anomenen EcoRI, arribant a la conclusió que tenen nivells de metilació molt alts, per sobre de la resta del genoma (Gama-Sosa et al. 1983b). Uns anys més tard, l'obtenció de seqüències corresponents a diferents famílies Alu va permetre veure que les diferents famílies estan molt metilades i que els elements Alu's més joves (els quals retenen més dinucleòtids CpG) podrien arribar a proporcionar fins a un 10% dels llocs de metilació del genoma (Schmid 1991).

La comparació de l'estat de metilació de les diferents famílies en diferents teixits va permetre veure nivells de metilació molt alts en teixits somàtics, però no en la línia germinal masculina, on els nivells més baixos de metilació correlacionen amb nivells detectables de trànscripció Alu (Hellmann-Blumberg et al. 1993; Kochanek et al. 1993). De forma similar, els elements repetitius LINE estan fortament metilats en la seva regió promotora, fet que provoca el silenciament dels dos trànscripcions per als quals codifica i que impedeixen la seva transposició (Furano et al. 1988; Nur et al. 1988; Woodcock et al. 1988). Les millores metodològiques introduïdes per l'aplicació de tècniques basades en la PCR han permès obtenir dades més precises que les obtingudes fins ara per mitjà de l'ús d'enzims de restricció i tècniques de hibridació (Weisenberger et al. 2005; Yang et al. 2004).

El fet que els elements repetitius acumulin nivells tant alts de metilació va fer pensar a alguns autors que potser les pèrdues de metilació globals que s'havien descrit en càncer podien tenir lloc de forma massiva en aquesta mena de seqüències, fet que ha quedat demostrat per diverses classes de seqüències repetitives com el DNA satèl·lit pericentromèric (Ji et al. 1997; Narayan et al. 1998; Qu et al. 1999; Weisenberger et al. 2005) i els elements dispersos Alu (Weisenberger et al. 2005; Yang et al. 2004) i LINE (Chalitchagorn et al. 2004; Choi et al. 2007; Estecio et al. 2007; Jurgens et al. 1996; Roman-Gomez et al. 2005; Weisenberger et al. 2005) (Revisat en (Ehrlich 2002)).

Tant per l'elevat nombre en què les trobem en el genoma com per la seva implicació en processos de desestabilització del genoma, les seqüències repetides Alu han estat objecte d'un profund anàlisi, més genètic que epigenètic, que ha evidenciat la seva importància en el modelatge dels genomes que les contenen així com la seva capacitat reguladora i inductora d'alteracions diverses, tant en càncer com en altres malalties. Vegem doncs què són les seqüències Alu i per què són importants en l'estudi del càncer.

Els elements Alu

Sota el nom de seqüències Alu hi trobem englobades diferents famílies de seqüències repetides relacionades entre sí, les quals prenen el seu nom de la diana de restricció de l'enzim AluI que contenen (Houck et al. 1979). La seva distribució queda restringida a primats, entre ells el genoma humà (Quentin 1994) on són l'element repetitiu que ha tingut un major èxit, arribant a existir més d'un milió de còpies per genoma haploide (Lander et al. 2001). Les seqüències Alu, conjuntament amb les seqüències repetides MIR, formen part d'una família més àmplia de repeticions anomenats SINEs, de l'anglès *Short Interspersed Nucleotide Element*.

Les seqüències Alu varen sorgir fa al voltant de 65 milions d'anys, en el període corresponent amb la radiació dels primats, per fusió de dues seqüències molt similars, el *free left alu monomer* (FLAM) i el *free right alu monomer* (FRAM) per a donar lloc a una estructura dimèrica d'uns 300 pb (Quentin 1992) (Figura 16 i 17). Tot i que existeixen petites diferències a nivell de seqüència entre el FLAM i el FRAM, tots dos provenen del 7SL RNA per deleció del seu domini S central (Revisat en (Batzer and Deininger 2002; Mighell et al. 1997)). El 7SL RNA codifica per

un RNA que forma part de la partícula de reconeixement del senyal o SRP, necessària per a la síntesi de proteïnes en el reticle endoplasmàtic (Walter and Blobel 1982).

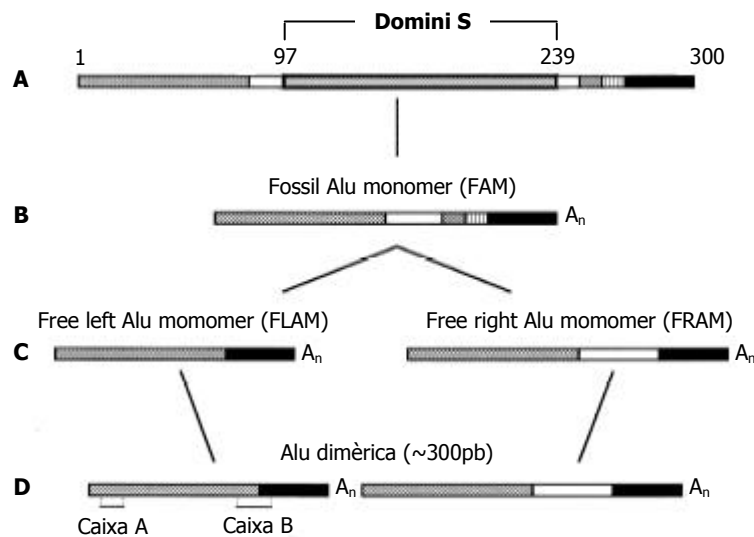


Figura 16. Origen de les seqüències Alu. A partir de la deleció del domini S del RNA 7SL (domini comprès entre el parell de bases 97 i 239) (A) va ser originat la seqüència ancestral FAM, la qual a més conté una cua d'adenines (A_n). Per divergència d'aquesta seqüència es varen originar el FLAM i el FRAM (C) els quals varen donar lloc, per fusió, de la primera molècula Alu dimèrica, la qual manté els dos tractes de poliadenina dels dos avantpassats FLAM i FRAM (D). Modificat de (Mighell et al. 1997).

Les seqüències Alu són retrotransposons que poden ser transcrits per la RNA polimerasa III, la qual reconeix les caixes A i B en la regió 5' de la seqüència (Figura 16 i 17). Això vol dir que el seu mecanisme de transposició inclou intermediaris de RNA. Donat que no codifiquen per cap proteïna o factor, es creu que obtenen els elements necessaris per a la seva transposició (retrotranscriptasa i endonucleasa) d'altres elements repetitius transposables com els element repetitius llargs dispersos o LINEs, els quals sí que codifiquen aquestes funcions (Revisat en (Batzer and Deininger 2002)). L'estat metilat de les seqüències Alu suprimeix la seva capacitat de transcriure's i transposar-se ja que les Alu's desmetilades són actives (Revisat en (Goll and Bestor 2005)).



Figura 17. Estructura genètica de la seqüència Alu. L'element Alu es troba entre dues repeticions directes (fletxes) les quals crea en el moment de la inserció. Les caixes A i B són necessàries per a la transcripció per RNA polimerasa III. L'element Alu no té terminador intern, així que la transcripció s'acaba quan la RNA pol III troba un terminador extern (TTTT). Els dos monòmers que conformen l'element Alu es troben units per un tracte poliadenina que presenta la seqüència consens A_5TACA_6 , a més a més de presentar una cua poliadenina final (AAA), la qual li és necessària per a la transposició. Modificat de (Batzer and Deininger 2002).

Tot plegat ha dut a alguns autors a suggerir un paper de defensa de la metilació enfront del perill que suposa l'activitat dels elements repetitius. Hi ha dues vies principals a través de les

quals els elements transposables i el DNA repetitiu en general poden alterar l'estructura del genoma. La primera via és a través de la disrupció gènica: en el procés de transposició, un element LINE o Alu pot insertar-se dins d'un gen i abolir-ne la funció. Aquesta mutagènesi per inserció està associada a la inactivació de gens causants de diferents malalties com la hemofília (inserció en el Factor IX), el síndrome d'Apert (Fgfr2), la deficiència de colinesterasa (colinesterasa), neurofibromatosi (Nf1) i leucèmia i càncer de mama (Brca1). A més a més, determinats elements Alu contenen seqüències de reconeixement per receptors d'algunes hormones esteroidees. D'aquesta manera, una Alu insertada a prop d'una regió promotora pot acabar per afectar l'expressió gènica (Revisat en (Batzer and Deininger 2002; Deininger and Batzer 1999)). La segona via implica un mecanisme que té efectes globals sobre el genoma. L'existència de al voltant de 1.1 milions de còpies d'Alu per genoma haploid (0.5 milions en el cas de LINEs) implica que poden actuar com a centres de recombinació homòloga al llarg de tot el genoma. Aquesta recombinació pot tenir múltiples conseqüències, com són les duplicacions, delecions, amplificacions i translocacions. Aquest segon mecanisme està també relacionat amb la generació de varies malalties com la α -talassèmia (α -globina), la distròfia muscular de Duchenne (DMD), trombofília (antitrombina), càncer de mama (Brca1), càncer de còlon HNPCC (Mlh1) i leucèmia mieloide crònica (MLL) (Revisat en (Batzer and Deininger 2002; Deininger and Batzer 1999)).

La desmetilació d'elements repetitius en càncer pot incrementar la taxa a la qual aquests dos mecanismes d'instabilització del genoma es donen en cèl·lules tumorals, per mitjà de la reactivació d'Alu's silenciades (increment de la transposició i mutagènesi) o a través de la descondensació de la cromatina i un increment de les taxes de recombinació entre seqüències repetides.

Conseqüències de la hipometilació global en càncer

Ja hem vist que la hipometilació global afecta bàsicament seqüències repetides. Les conseqüències d'aquesta hipometilació han estat, i encara són, objecte de debat.

Una de les primeres pistes va venir de l'estudi d'un síndrome anomenat Immunodeficiència, inestabilitat centromèrica i anomalies facials o ICF. Aquest és un síndrome genètic recessiu rar el qual es caracteritza, entre d'altres, per presentar un alt nombre de reorganitzacions cromosòmiques que afecten sobre tot a les regions pericentromèriques dels cromosomes 1, 16 i 9 (Revisat en (Ehrlich 2002)). Aquestes regions pericentromèriques estan constituïdes bàsicament per DNA repetitiu tipus satèl·lit, el qual està fortament metilat en cèl·lules normals. Curiosament, limfòcits de pacients de ICF pateixen una hipometilació severa d'aquest DNA repetitiu pericentromèric en els mateixos cromosomes en els que es donen les reorganitzacions (Jeanpierre et al. 1993; Miniou et al. 1994). Estudis genètics en pacients amb ICF varen determinar que són mutacions en el gen de la Dnmt3b les responsables d'aquest síndrome (Hansen et al. 1999; Okano et al. 1999; Xu et al. 1999). Aquests resultats suggerien un paper de la hipometilació com a inductor d'instabilitat cromosòmica i per tant, començaven a indicar quina podia ser la relació entre hipometilació i inestabilitat en càncer.

De forma similar, la inducció farmacològica de desmetilació en cèl·lules de mamífer amb l'anàleg de base 5-aza-2'-deoxicitidina, té uns efectes similars als vistos en pacients amb ICF, els quals inclouen la desmetilació i descondensació de regions riques en DNA satèl·lit pericentromèric en els cromosomes 1, 9 i 16. Aquesta descondensació s'associa amb un increment en els intercanvis de cromàtides germanes (per mitjà d'un increment de la recombinació mitòtica) i les endoreduplicacions (duplicació de la càrrega genètica complerta per mitjà d'una divisió cel·lular sense cariocinesi) (Hori 1983; Stopper et al. 1992). Recentment ha estat descrit que la hipometilació de regions subteloemèriques incrementa la longitud i la inestabilitat de les regions telomèriques adjacents, provocant un increment de la recombinació homòloga (Gonzalo et al. 2006). Aquests resultats no fan sinó sustentar la hipòtesi que la metilació del DNA juga un paper en el manteniment de l'estructura de la cromatina i la supressió de la recombinació.

En conjunt, totes aquestes evidències apuntaven a un paper inestabilitzador i per tant tumorigènic de la hipometilació. Una visió oposada dels efectes de la hipometilació en càncer va ser la plantejada pel grup de Rudolf Jaenisch treballant amb un model de càncer murí, els ratolins APC Min. Aquests ratolins tenen alterada genèticament una còpia del gen APC, model que mimetitzava el síndrome de poliposi hereditària i que predisposa als ratolins a patir poliposi intestinal. En aquest model, la inducció d'hipometilació per reducció de l'activitat metiltransferasa i el subministrament de baixes dosis de la droga desmetilant 5-aza-deoxicitidina provocava una reducció en el nombre de pòlips intestinals (Laird et al. 1995). Les conclusions d'aquest treball, tot i que restringides al model usat, no fan sinó incrementar la confusió al voltant del paper de la hipometilació en càncer.

Posteriorment, relativament pocs treballs han abordat de forma directa les causes i conseqüències de la hipometilació en càncer. En aquest context cal ressaltar els treballs del mateix Jaenisch i col·laboradors en ratolí, els quals han demostrat un paper causal de la hipometilació en la generació d'inestabilitat i càncer, tant en models *in vitro* com *in vivo*. L'obtenció de ESC murines deficientes per Dnmt1 (Li et al. 1992) va permetre a aquest grup estudiar l'efecte de la hipometilació sobre l'estabilitat del genoma. Així doncs varen poder demostrar que la hipometilació global provocada per l'absència d'activitat Dnmt1 en aquestes cèl·lules, comportava un increment de 10 vegades de la taxa de mutació en el locus lligat al cromosoma X Hprt i de 6 vegades sobre la taxa de mutació del gen Tk inserat en un autosoma. En ambdós casos, les mutacions eren degudes bàsicament a alteracions estructurals del tipus insercions i delecions (Chen et al. 1998). De forma similar, treballs realitzats per Rizwana i col·laboradors demostraven en línies tumorals murines que la presència d'illes CpG desmetilades en petits fragments de DNA extracromosòmics coneguts com a dobles diminuts, incrementava la taxa de dimerització i delecions dels dobles diminuts, respecte a dobles diminuts amb les illes metilades (Rizwana and Hahn 1999). Ja havíem vist que la disrupció en homozigosi de Dnmt1 en ratolins és letal (Li et al. 1992), de manera que no existia cap model d'estudi de la hipometilació *in vivo*. L'any 2003 es publicava un treball realitzat en el grup de Jaenisch que reportava l'obtenció de ratolins amb una severa reducció de l'activitat de Dnmt1, fet que provocava una profunda hipometilació i l'aparició de limfomes de cèl·lules T molt agressius entre els 4 i els 8 mesos de vida (Gaudet et al. 2003). De forma significativa, els

limfomes provocats en aquests ratolins tenien múltiples reorganitzacions cromosòmiques, tant a nivell de guanys com de pèrdues de cromosomes sencers o fragment cromosòmics. Utilitzant un procediment experimental similar, varen introduir ESC amb la reducció de Dnmt1 que havien usat prèviament, en un altre model murí amb mutacions en heterozigosi en els gens Nf1 i p53. La presència d'aquestes mutacions predisposa a aquests ratolins a una major incidència de sarcomes, per pèrdua de l'altra còpia activa de Nf i p53. Aquesta major incidència es veia encara més incrementada en els ratolins amb la deficiència de Dnmt1 (Eden et al. 2003). De forma similar, els tumors generats mostren un increment en la taxa de LOH en els loci de Nf i p53.

La manca de models de càncer humans va dur a alguns investigadors a realitzar la deleció de gens de metiltransferases en la línia cel·lular de CCR cromosòmicament estable HCT116. Sorprenentment, només la doble deleció Dnmt1 i Dnmt3b (Rhee et al. 2002), i no la de Dnmt1 sola (Rhee et al. 2000) reduïa considerablement els nivells de metilació. La reducció en els nivells de 5mC en la HCT116 amb la doble deleció correlacionaven amb una major incidència d'alteracions cromosòmiques estructurals (Karpf and Matsui 2005). Experiments posteriors duts a terme sobre la mateixa línia HCT116 han demostrat però que la inactivació de Dnmt1 per mitjà de RNA interferent sí que té efectes sobre els nivells de metilació (Robert et al. 2003). L'aparent contradicció entre ambdós resultats prové del fet que la supressió de l'activitat de Dnmt1 per mitjà de recombinació homòloga que varen dur a terme Rhee i col·laboradors no va aconseguir eliminar les dues còpies del gen Dnmt1.

Aquests experiments demostren que almenys en determinats models experimentals la hipometilació global induïda té un paper causal sobre la inestabilitat genètica, fet que suggereix, però no demostra, que aquesta situació es pugui donar també en càncer.

La hipermetilació de illes CpG en tumors

Aquest és el mecanisme més conegut sobre com la metilació del DNA afecta al procés tumoral. Des del moment en què es va descobrir el primer gen hipermetilat en càncer (Baylin et al. 1987) el nombre de gens silenciats per aquest mecanisme no ha parat de créixer (Taula 8), i quasi podríem dir que aquest nombre supera al dels gens mutats en càncer (Revisat en (Jones and Baylin 2002)). Aquest mecanisme involucra gens en totes les localitzacions cromosòmiques i en virtualment totes les vies cel·lulars conegudes. Els gens silenciats per aquest mecanisme els podem agrupar en tres classes. En primer lloc aquells gens supressors tumorals clàssics que ja s'han trobat mutats en la línia germinal de famílies amb síndromes hereditaris de càncer. En segon lloc gens candidats supressors tumorals. I en tercer lloc, gens trobats en diferents *screenings*, la funció dels quals es desconeix (Revisat en (Baylin 2005; Ehrlich 2002; Feinberg and Tycko 2004; Herman and Baylin 2003; Jones and Baylin 2002; Jones and Baylin 2007; Jones and Laird 1999)).

Taula 8. Gens hipermetilats en càncer.

Gen	Funció	Tipus de tumor	Conseqüències
Mlh1	DNA <i>Mismatch repair</i>	Còlon, endometri, estómac	Mutacions <i>frameshift</i>
Brca1	Reparació DNA, transcripció	Mama, ovari	Trencaments doble cadena?
p16 ^{INK4a}	Inhibidor de kinasa depenent de ciclina	Múltiples tipus	Entrada en cicle
p14 ^{ARF}	Inhibidor de MDM2	Còlon, estómac, ronyó	Degradació de p53
p15 ^{INK4b}	Inhibidor de kinasa depenent de ciclina	Leucèmia	Entrada en cicle
MGMT	Reparació de O6 alquil guanina	Múltiples tipus	Mutacions, quimiosensibilitat
GSTP1	Conjugació a glutatió	Pròstata, mama, ronyó	Acumulació d'adductes?
p73	Homòleg de p53	Limfoma	Desconegut
LKB1	Serina- Treonina kinasa	Còlon, mama, pulmó	Desconegut
ER	Receptor d'estrògens	Mama	Insensibilitat hormonal
PR	Receptor de progesterona	Mama	Insensibilitat hormonal
AR	Receptor d'andrògens	Pròstata	Insensibilitat hormonal
PRLR	Receptor de prolactina	Mama	Insensibilitat hormonal
TSHR	Receptor d'hormona estimulant de tiroides	Tiroides	Insensibilitat hormonal
RAR β 2	Receptor d'àcid retinoic β 2	Còlon, pulmó, cap i coll	Insensibilitat a vitamines?
CRBP1	Proteïna d'unió a retinol	Còlon, estómac, limfoma	Insensibilitat a vitamines?
RASSF1A	Homòleg de l'efector de Ras	Múltiples tipus	Desconegut
NORE1A	Homòleg de l'efector de Ras	Pulmó	Desconegut
VHL	Component ubiquitin lligasa	Ronyó, hemangioblastoma	Pèrdua resposta a hipòxia?
Rb	Inhibidor de cicle cel·lular	Retinoblastoma	Entrada a cicle
THBS1	Trombospondina 1. Antiangiogènica	Glioma	Neovascularització
CDH1	E-cadherina. Adhesió cel·lular	Mama, estómac, leucèmia	Disseminació
CDH13	H-cadherina. Adhesió cel·lular	Mama, pulmó	Disseminació?
FAT	Cadherina. Supressor tumoral	Còlon	Disseminació?
HIC1	Factor transcripció	Múltiples tipus	Desconegut
APC	Inhibidor de β -catenina	Tracte aerodigestiu	Activació via β -catenina
SFRP1	Proteïna secretada relacionada amb <i>frizzled</i>	Còlon	Activació via Wnt
DKK1	Inhibidor de Wnt extracel·lular	Còlon	Activació via Wnt
WIF1	Factor inhibidor de Wnt	Còlon, pulmó	Activació via Wnt
COX2	Ciclooxygenasa 2	Còlon, estómac	Resistència antiinflamatòria?
SOCS1	Inhibidor de la via JAK-STAT	Fetge, mieloma	Activació de JAK2
SOCS3	Inhibidor de la via JAK-STAT	Pulmó	Activació de JAK2
GATA4	Factor transcripció	Còlon, estómac	Silenciament de gens diana
GATA5	Factor transcripció	Còlon, estómac	Silenciament de gens diana
ID4	Factor transcripció	Leucèmia, estómac	Desconegut
SRBC	Proteïna d'unió a BRCA1	Mama, pulmó	Desconegut
SYK	Tirosina kinasa	Mama	Desconegut
RIZ1	Histona/proteïna metiltransferasa	Mama, fetge	Expressió gènica aberrant?
DAPK	Pro apoptòtic	Limfoma, pulmó, còlon	Resistència a apoptosi
TMS1	Pro apoptòtic	Mama	Resistència a apoptosi
IGFBP3	Proteïna d'unió a factor de creixement	Pulmó, pell	Resistència a apoptosi
TPEF	Proteïna transmembrana	Còlon, bufeta	Desconegut
SLC5A8	Transportador de sodi	Glioma, còlon	Desconegut
HOXA9	Proteïna homeobox	Neuroblastoma	Desconegut
EXT1	Síntesi d'heparan sulfat	Leucèmia, pell	Trencament unions cel·lulars
Lamina A/C	Filament nuclear intermedi	Limfoma, leucèmia	Desconegut
WRN	Reparació DNA	Còlon, estómac, sarcoma	Trencaments DNA

Taula modificada de (Esteller 2007).

El gran nombre de gens hipermetilats i silenciats en càncer (entre 100 i 400, depenent de l'aproximació usada i el tipus de tumor; revisat en (Esteller 2007)) ha dut a alguns autors a hipotetitzar l'existència d'un fenotip metilador o CIMP (*CpG island methylator phenotype*) (Toyota et al. 1999a). Aquest suposat fenotip es defineix per analogia al fenotip mutador en el cas dels tumors colorectals amb inestabilitat de microsatèl·lits. En aquest últim cas, la presència massiva de mutacions ve provocada per alteracions genètiques en els gens del *mismatch repair* (Revisat en (Kinzler and Vogelstein 1996)). De forma similar, els partidaris de l'existència d'un fenotip metilador en càncer argumenten que el gran nombre de gens hipermetilats en càncer obeeix a la pertorbació d'algun sistema cel·lular de control d'aquests mecanismes epigenètics. El fet que aquest fenotip no quedi ben definit, així com la manca d'evidències que demostrin que es provocat per alteracions en alguna via concreta han fet que la seva existència quedi en entredit (Yamashita et al. 2003). El descobriment que molts dels pacients de HNPCC amb inestabilitat de microsatèl·lits tenen el gen MHL1 silenciats per metilació (Herman et al. 1998), no va fer més que incrementar la polèmica entorn d'aquest tema. Per a alguns autors, una hipotètica inestabilitat epigenètica seria la responsable de la metilació, entre d'altres gens, de MLH1, fet que conduiria a les cèl·lules a una inestabilitat genètica afegida. Així doncs, en aquest model la inestabilitat epigenètica precediria a la inestabilitat genètica en aquests pacients de càncer MIN positius.

La metilació de les illes CpG sempre s'ha presentat com un fenomen que afecta a gens discrets. Això és, si bé la hipometilació és un fenomen global que pot afectar a regions més o menys grans del genoma, la hipermetilació del DNA sempre s'ha descrit com un fenomen que afecta a gens aïllats en diferents punts del genoma. Treballs realitzats per membres del nostre grup en col·laboració amb Susan Clarck han fet canviar aquest paradigma, tot demostrant que la metilació del DNA es pot estendre sobre regions cromosòmiques extenses, afectant les illes CpG de múltiples gens veïns (Frigola et al. 2006). En aquest treball, es descriu la hipermetilació de tres grans blocs de DNA contigus en la regió cromosòmica 2q14.2, el més gran dels quals engloba la metilació de 7 illes CpG al llarg de 1 Mb. Múltiples preguntes es deriven d'aquest treball: per quins mecanismes alguns gens es metilen i d'altres no? És a dir, què dirigeix la metilació del DNA? Quina relació s'estableix entre les modificacions de les histones i la metilació del DNA? Hi ha altres regions com aquesta en CCR o d'altres tipus de tumors? Certament, queda molt treball i molts experiments per a poder donar resposta a aquestes i d'altres preguntes.

5. Tècniques d'anàlisi de la metilació

El gran avenç en la identificació i estudi d'alteracions de la metilació en tumors ha vingut de la mà del desenvolupament d'un extens catàleg de metodologies disponibles. L'estudi de la metilació es pot fer des de molts punt de vista, i per tant la metodologia a emprar dependrà de la pregunta que ens formulem (Revisat en (Cottrell 2004)). A continuació veurem les principals metodologies usades tenint en compte les seves principals aplicacions.

Quantificació de nivells globals de 5mC

Una de les primeres preguntes que ens podem plantejar és si un determinat genoma conté 5mC o no. Les tècniques freqüentment usades es basen en el fet que la 5mC té una mobilitat cromatogràfica o electroforètica diferent de la de la citosina i la resta de bases. Les dues tècniques més emprades s'anomenen *High-performance liquid chromatography* (HPLC) i *High-performance capillary electrophoresis* (HPCE), i totes dues mostren nivells de sensibilitat molt alts en la detecció de 5mC, tot i que no permeten determinar-ne la distribució (Revisat en (Esteller 2007)).

Estudis globals dels patrons de metilació i identificació de noves dianes

Les metodologies que veurem en aquest apartat es basen en la utilització de tres tècniques fonamentals: els enzims de restricció, el tractament de DNA amb bisulfit sòdic i la PCR.

Els enzims de restricció són proteïnes d'origen bacterià que tallen doble cadena de DNA quan troben la seva seqüència de reconeixement i tenen el seu origen com a mecanisme de defensa bacterià enfront de DNA exogen. Els més àmpliament utilitzats com a eina en biologia molecular són els enzims de restricció de classe II, els quals tallen dins de la seva seqüència de reconeixement sempre de la mateixa manera amb gran precisió. L'existència d'enzims de restricció de classe II sensibles i insensibles a la presència de 5mC dins de la seqüència de reconeixement va fer que ràpidament fossin incorporats en l'anàlisi de patrons de metilació. L'ús d'enzims de restricció limita l'anàlisi a un o uns pocs residus de citosina per cada regió, fet que queda compensat amb la possibilitat d'analitzar múltiples regions de forma simultània.

Paral·lelament, una de les tècniques que més ha revolucionat l'estudi de la metilació ha estat el tractament de DNA amb bisulfit sòdic (Figura 18) (Clark et al. 1994). Aquesta tècnica basa el seu principi en la conversió de la citosina a uracil depenent del seu estat de metilació: si bé la citosina no metilada es transforma a uracil, la 5mC resta com a tal després de la reacció. La seqüenciació del DNA tractat determinarà quines citosines estaven metilades (es seqüencien com a citosina) i quines estaven no metilades (es seqüencien com a timina). Un dels principals avantatges d'aquesta tècnica és la possibilitat d'analitzar múltiples citosines en una mateixa regió, al màxim nivell de resolució.

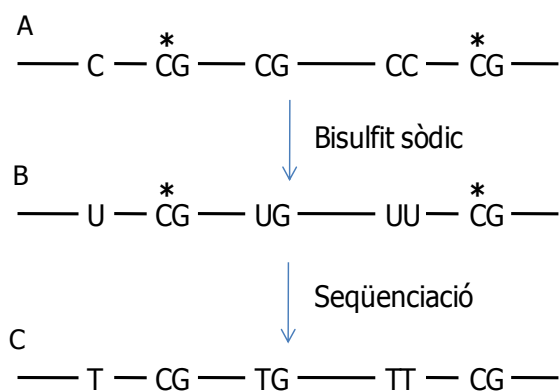


Figura 18. Tractament de DNA amb bisulfit sòdic. El DNA, representat per la línia horitzontal, conté tant citosines en posició 5' a una guanina (CG) com en posició 5' a qualsevol altre nucleòtid (C). L'asterisc sobre determinades citosines n'indica l'estat metilat (A). Quan tractem el DNA amb bisulfit sòdic, totes les citosines no metilades són convertides a uracil (U) mentre que les citosines metilades romanen sense modificar (B). La comparació entre el DNA seqüenciat i la seqüència original ens permet identificar quines posicions estaven metilades i quines no (C).

Finalment, el gran desenvolupament de les aplicacions de la PCR a la biologia molecular han fet que avui en dia disposem d'un ampli ventall de mètodes basats en diferents combinacions de les tres tècniques (enzims de restricció, bisulfit sòdic i PCR, figura 19) que permeten realitzar des dels anàlisis més específics als més globals. Si bé la PCR ens permet treballar amb quantitats molt petites de mostra (donat que aquesta serà posteriorment amplificada), l'amplificació preferencial de determinades seqüències pot donar lloc a resultats esbiaixats. No obstant, en determinades ocasions es cerca esbiaixar de forma deliberada l'amplificació per tal d'analitzar un determinat compartiment genòmic. La figura 19 mostra de forma esquemàtica les principals metodologies usades en l'anàlisi de la metilació i les tècniques moleculars en les que es basen. Podem prendre bàsicament dues rutes a partir d'un DNA purificat.

En primer lloc podem començar per digerir el DNA genòmic amb enzims de restricció. Entre les metodologies que segueixen aquesta via, hi trobem el *Restriction Landmark Genomic Scanning* (RLGS) (Hayashizaki et al. 1993) (Figura 19A). Degut a que RLGS no usa cap pas d'amplificació n'evita els possibles artefactes de la PCR, però requereix l'ús de grans quantitats de mostra, fet no sempre possible quan treballem amb mostres clíniques com els tumors primaris. Si disposem de poca quantitat de mostra, el més adequat és usar tècniques que combinin l'ús de la PCR. Diferents metodologies d'anàlisi global segueixen aquest esquema. Les principals diferències entre elles les trobem en els passos d'amplificació i el mètode per mostrar i analitzar els productes amplificats. En primer lloc podem esmentar l'adaptació a l'estudi de la metilació de la tècnica *Amplified fragment length polymorphism* (AFLP), la qual passa a anomenar-se *Methylation-sensitive amplified fragment length polymorphism* (MS-AFLP) (Vos et al. 1995)(Figura 19B). Per mitjà de l'ús d'enzims de restricció sensibles i insensible a la metilació, la posterior lligació d'adaptadors de seqüència coneguda i l'amplificació per PCR es generen patrons complexos de bandes que poden ser visualitzats fàcilment en gels de poliacrilamida desnaturalitzants d'alta resolució. D'altra banda, la tècnica *Methyl CpG Amplification* (MCA) usa un esquema similar de generació dels fragments amplificats, tot i que usa la hibridació amb sondes per a determinar si una seqüència coneguda està o no metilada i permet el descobriment de noves dianes per mitjà de la tècnica *Representational difference analysis* (RDA)(Toyota et al. 1999b) (Figura 19C). Basant-se en MCA, la tècnica *Amplification of intermethylated sites* (AIMS) simplifica enormement la identificació i el recompte d'alteracions de la metilació de MCA a través de la generació de *fingerprints* de complexitat variable amb gels de poliacrilamida desnaturalitzants d'una forma similar als patrons obtinguts per MS-AFLP (Frigola et al. 2002) (Figura 19B).

Si el nostre punt de partida és el tractament amb bisulfit sòdic del DNA purificat, disposem d'un ventall de metodologies bàsicament adreçades a la validació i l'estudi específic de dianes de metilació. Un cop el DNA ha estat tractat amb bisulfit podem realitzar la seqüenciació directa d'una determinada regió (Figura 19D). Si prèviament a la seqüenciació clonem els fragments amplificats el que farem serà un anàlisi d'al·lels, és a dir, per mitjà de la clonació analitzem únicament una monocadena de DNA provinent d'una única cèl·lula, mentre que si realitzem seqüenciació directa analitzarem un conjunt d'al·lels provinents de cèl·lules diferents. Ambdós mètodes permeten l'anàlisi de múltiples dinucleòtids CpG a la vegada.

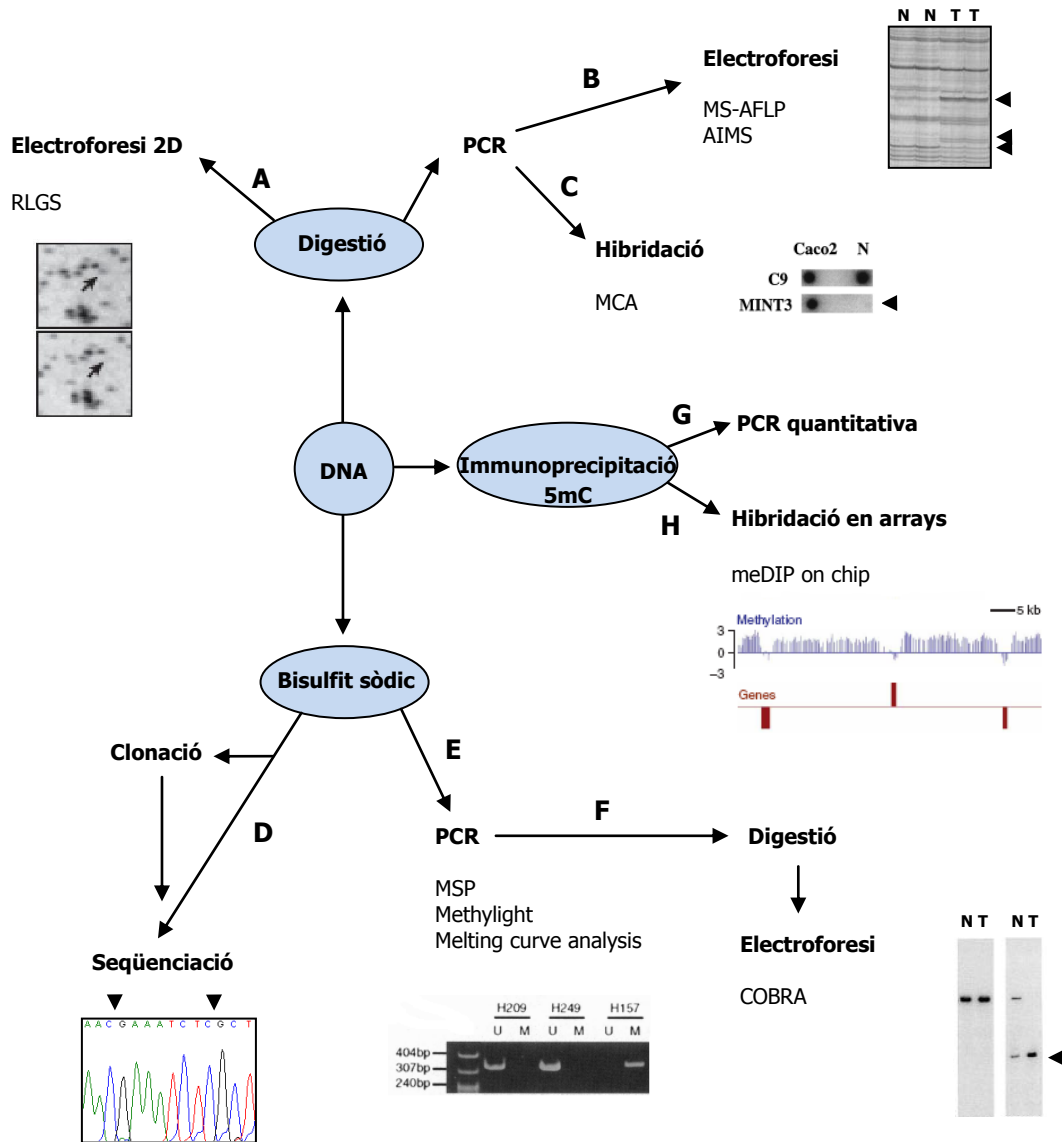


Figura 19. Metodologies d'anàlisi de la metilació del DNA. Aquest diagrama esquematitza les diferents combinacions tècniques que podem usar en l'anàlisi de la metilació del DNA. La desaparició d'un punt (fletxa) en un gel 2D de RLGS és indicatiu d'hipermetilació (A). En el cas d'AIMS, els productes amplificats són visualitzats en gels desnaturalitzants, en els quals podem veure diferències (fletxes) entre els patrons originats de teixit normal (N) i tumoral (T)(B). La tècnica MCA permet la visualització de metilació diferencial a través de la hibridació Dot Blot. En la figura veiem absència d'hibridació del clon MINT3 (fletxa) en teixit normal (N) però no en la línia de CCR CaCo2 (C). La seqüenciació directa o prèvia clonació de DNA tractat amb bisulfit ens permet conèixer quines citosines estan metilades (es llegeixen com a G en la seqüència, fletxes) i quines no (llegides com a T). L'amplificació de DNA tractat amb bisulfit amb primers específics de producte metilat (M) i no metilat (U) ens permet veure en un gel d'agarosa l'estat de metilació d'una CpG en concret per mitjà de MSP. En aquest cas, els tumors H209 i H249 amplifiquen únicament amb el set de primers específics de producte no metilat mentre que el tumor H157 amplifica únicament amb els primer de producte metilat (E). La tècnica COBRA aprofita els canvis de seqüència introduïts pel tractament amb bisulfit sòdic. Així, la presència de citosines no metilades dins la seqüència de reconeixement de l'enzim de restricció provoquen la destrucció de la diana després del tractament fet que comporta la no digestió del producte amplificat. En canvi si la citosina en la diana està metilada la diana es manté i l'enzim talla el producte amplificat, fet que comporta l'aparició de bandes de menor tamany (fletxa)(F). Finalment, la immunoprecipitació de DNA amb anticossos contra 5mC (meDIP) pot ser quantificada per mitjà de PCR a temps real (G) o bé mitjançant la hibridació sobre arrays (H). La imatge mostra com els nivells de metilació (barres blaves) són més baixos en regions intragèniques que en regions intergèniques (H).

Donat que el tractament amb bisulfit sòdic modifica la seqüència del DNA, podem usar diferents estratègies per tal de discriminar entre el producte metilat i el no metilat, per mitjà de l'ús d'amplificació per PCR i el tall amb enzims de restricció.

La primera estratègia consisteix en amplificar de forma directa DNA tractat amb bisulfit sòdic utilitzant primers dissenyats específicament per anellar amb producte metilat (conserva les citosines en CpG) i producte no metilat (les citosines en CpG es converteixen en timines). Aquest mètode d'anàlisi específic s'anomena *Methyl-specific PCR* (MSP) (Herman et al. 1996) (Figura 19E). Donada la seva alta sensibilitat ha de ser posat a punt acuradament per tal d'obtenir una bona especificitat. Per aquest motiu és essencial disposar de controls tant de producte metilat com no metilat per tal d'evitar falsos positius i negatius. Aquest és un mètode ràpid i senzill adreçat bàsicament a la detecció de marcadors de metilació en mostres clíniques. Algunes variacions d'aquesta tècnica inclouen la detecció dels productes amplificats per mitjà de sondes de hibridació amb PCR a temps real (Methylight)(Eads et al. 2000)(Figura 19E). Per mitjà de la tecnologia de la PCR a temps real també podem realitzar un anàlisi de corbes de fusió o *melting curve analysis* (Worm et al. 2001)(Figura 19E). Aquest mètode és basa en el fet que una seqüència metilada conserva les citosines després de ser tractada amb bisulfit sòdic de manera que la seva temperatura de fusió és més elevada que una seqüència no metilada. Per tant dos productes amplificats provinents d'una mateixa regió genòmica diferencialment metilats, presentaran típicament una diferència de temperatura de fusió determinada per a cada seqüència, essent aquesta temperatura més elevada per a la seqüència metilada. Aquest fet s'explica per la major estabilitat dels ponts establerts entre citosina i guanina que els establerts entre timina i adenina.

Una segona estratègia consisteix en el tall del producte amplificat amb enzims de restricció que continguin CpG en la seva seqüència diana. Si el producte de PCR prové d'una seqüència lliure de metilació, la citosina es converteix en timina i la diana es perd, de manera que l'enzim no talla. En canvi si la citosina en la diana estava metilada, la citosina no es converteix i la diana es conserva, permetent el tall amb l'enzim de restricció. Aquest mètode s'anomena *Combined bisulfite restriction analysis* (COBRA) (Xiong and Laird 1997)(Figura 19F).

En darrer lloc, la disponibilitat d'anticossos que reconeguin 5mC en DNA purificat ha obert les portes a l'anàlisi dels patrons de metilació genòmics d'alta resolució per mitjà de la hibridació dels productes immunoprecipitats sobre *arrays* de DNA d'alta densitat (Figura 19H). Aquesta metodologia (meDIP on chip), tot i que proporciona un nombre de dades elevadíssim, requereix d'un gran esforç d'anàlisi computacional de les mateixes. D'altra banda, ens permet obtenir una imatge global molt precisa dels patrons de metilació d'una forma no esbiaixada, en el cas que s'utilitzin *arrays* genòmics. Alternativament, podem centrar-nos en l'estudi de regions concretes per mitjà de la quantificació dels productes immunoprecipitats amb PCR a temps real (Figura 19G).

Per a una revisió molt més extensa dels mètodes d'anàlisi de la metilació referir-se a (Varis autors 2005).