Development of model systems to reconstruct the unicellular prehistory of animals: an emphasis on the cell cycle

Alberto Perez-Posada

TESI DOCTORAL UPF / 2019

Thesis supervisor

Dr. Iñaki Ruiz-Trillo

Institute of Evolutionary Biology (CSIC-UPF)

DEPARTAMENT DE CIÈNCIES EXPERIMENTALS I DE LA SALUT

A mi padre, por llevarme en bicicleta por el Monte de la Rana.
A mi madre, por los paquetitos de animales y dinosaurios.
A mis hermanos, por haber conquistado la casa y cambiarlo todo.

A ellos cuatro, por haberme apoyado
incondicionalmente en este viaje.

En Barcelona, a día 26 de Septiembre de 2019

Acknowledgements

(léase con acento andaluz)

Los que me conocen saben bien que me olvido de muchas cosas, pero guardo en el recuerdo los detalles más pequeños e inesperados. Por eso es por lo que, para mí, este viaje comenzó hace muchísimo tiempo, y hay muchos nombres a los que estar agradecido. La mayoría me habéis enseñado cosas valiosísimas y me habéis ayudado en un sinfín de momentos, y sin vosotros el camino habría sido imposible o ni siquiera hubiera existido, como bien saben algunos.

Primero de todo, gracias a **Iñaki**, por haberme abierto las puertas a su lab y haberme permitido formar parte de la gran familia que es el MCG, brindándome una oportunidad y recursos inigualables para aprender a hacer ciencia. O, dicho de otro modo, gracias por haberme respondido a aquel email desorbitadísimo en 4º de carrera pidiéndote hacer la tesis, por haber sabido ver en mí el entusiasmo que me llevé en la mochila a esa entrevista exprés en Barcelona, y por haber sabido escucharme, calmarme, dejar que me equivoque y que arregle mis errores, y animarme en los momentos más duros. Pero, sobre todo, gracias por haber vuelto a recibirme con los brazos abiertos en esta última etapa. Siento que lo mejor que me llevo de este tiempo no es sólo lecciones de cómo llevar un lab: también me llevo un mentor, un colega, y la sensación de tener muchas cosas buenas en común con alguien a quien admiro mucho.

Gracias a todo el elenco de grandes personas que hacéis del MCG lab una gran familia:

Muchísimas gracias a **Helena** por haberme acogido en tu poyata, por haberme enseñado tanto desde el primer día, por las transfectoparties, los *Wikipedia times*, los buenos y malos momentos juntos, por ser una super compi de piso, y por ese aire tolkieniano que le das a las cosas. Eres una campeona y uno de mis ejemplos a seguir, ¡estoy muy orgulloso de haber sido entrenado por ti! **Núria**, muchísimas gracias por tu ayuda incondicional en todas mis propuestas, tanto conjuntas como en mis propios proyectos; gracias por el cuquismo, las discusiones científicas, los journal clubs, las teorías y misterios de la ciencia ficción, las charlas volviendo del lab, y gracias por tu sonrisa, tus abrazos y el buen humor cada día sin falta. Muchas gracias a **Sebas**, por ser otro de mis grandes ejemplos y haberse ganado el alta estima que tengo por él. Gracias por tu humor y tus bromas, por una muy necesitada sesión de arcoterapia en

2018, y por tus historias de todo tipo: ya sean científicas, ya sean surrealistas, pero siempre muy valiosas para mí. Gracias también a tu familia, a **Yani** y a **Mirko**, por haberme mostrado tanto cariño. Mis mejores deseos para Najle Jr; da igual que pierda la paciencia al colorear y se salga, ¡a mí también me pasaba!

Gracias a **Edu**, mi compañero de aventuras durante toda la tesis. Me he visto reflejado en ti muchas veces, y otras tantas me ha salvado poder hablar contigo de las cosas. Por haber compartido tantos momentos juntos, tantas conversaciones acerca del PhD, de nuestros objetivos en la vida, de las inseguridades y aspiraciones como científicos. Por las filogenias de las CDKs, Ezekiel, el DMSO, el esplendor geométrico, los tributos a Oasis, y el apoyo mutuo incondicional. Dice el Kunkeiro69 que es strongly likely que vengamos de la ameba, Edu. Definitely Maybe. A **Alicia**, por ser como una hermana. Por todas las miradas de "¿qué puñetas estamos haciendo con nuestras vidas?", la risa floja y tonta, los acentos inventaos, los post-its con flamencas, las naranjas (¡¡las naranjas!!), el cuqui committee seal of approval en todas mis figuras, las dudas de R que se convertían en charlas súper personales y profundas que salían de la nada, tu descubrimiento de la música electrónica, ¡y tó ese poderío que llevas dentro, rubia!

Muy importante para mi proyecto ha sido el trabajo conjunto con **Andrej** y **Omaya**; con ellos pulí el espíritu crítico ante la ciencia y el pragmatismo al resolver los problemas, y la infinidad de discusiones y puestas en común de seguro me han curtido para defender cualquier proyecto. Thanks **Andrej** for the many experiences together, for all the critical thinking and your knowledge, for the trip to Japan and, curiously enough, for indirectly helping me in reassuring who I truly am and who I want to be. Thanks, **Omaya**, for your mentoring, your microscopy, the moments of absurd humor and surreal faces, your pinch of za'atar, your cocktails, the D__E__E__P House Music, and many more things, and because this time the answer is YES.

A **Meri** por su infinita paciencia con mis primers perdidos, su apoyo en los experimentos, por todos los momentos surrealistas y sus recuerdos de Cái, por buscarme cuando "me perdía", por sus cancionces de trap chungas, y por básicamente mirarme desde su silla y seguirme la corriente mientras me da el sirocazo ocasional. Gracias a mis coetáneas **Ola** y **Kons** por charlas desde lo más científico a lo más abstracto a lo más místico a lo más absurdo; a **Dan** (por Jane, tus plantas, tu insight en los análisis y overall por ser tan molón), a **Michelle** (por tu sabiduría, tu buen corazón, tu empatía, tu humor exquisito y overall por ser tan molona). Hemos sido tantos que me da cosa dejarme a alguien…

Muchas gracias y recuerdos a todos los que han pasado por el MCG: los que ya marcharon, como **David** (por tu buen rollo y tu Deivis-chill spirit contagioso), **Xavi Grau** (por recordar mi llegada al lab como un epítome del romanticismo, ¡y por tus PCAs preciosos!), **Xavi Florenza** (por todos los recuerdos del máster, las becas, los sándwiches de la máquina, y por haber ganado: jamás encontré la pegatina de VISITANT), **Mati** (te juro que en mi vida me he reído en un lab como con tus cosas); y a los que llegaron hace menos (**Jesús**, **Ona**, **Mami**), y a todos los demás.

No puedo acabar sin mencionar Japón. Gracias a **Hiroshi** (que puede leer español) por abrir las puertas de su lab y hacer de Japón una experiencia inolvidable. Dada mi historia romántica con el lab, conocerte en persona fue un hito personal para mí. Además, como ya dije a muchos, "quedaos con todo Tokyo y su turismo, y dejadme a mí los campos, la gente, el sushi y el donburi tirao de precio". Gracias a **Takaaki** (Burosa!), **Arisa**, **Kazusa**, **Kaho**, **Mitsu**, **Katsu**, **Seitaro**, **Sung**, **Aizawa-sensei**, **Pablo**, y a todos los que lo hicieron posible. 私は永遠に感謝しています. ヘンタイケソ! Blessing you with happy futures; so, see you again.

La tesis habría sido muy diferente de no tener por amigos a la panda de cafres también conocida como "becarios de la Caixa™®". A **Patxi**, por haber roto el hielo el primer año con aquel kebab en Zaragoza, poniendo velitas a los santos, las peleas de gallos arqueólogos, historias de perros y hámsters, el chiste del elefante y el conejo, y charlas vitales varias; a **Juanma**, por hacer junto con **Simón** el otro dúo gallego-andaluz, habernos obsequiado con el Cocodrilo y hacer, junto con mi paisano **Pablo** y con **Ana** y su espíritu de feria, que se note la fuerza andaluza. A **Simón** y a **Ander** por los memes, los memes de memes, los steamed hams, el work-life balance, los ánimos a nivel personal y las cosas de ingenieros en general. A **Lidia**, por otras tantas charlas y ánimos y sonrisas. A **Víctor**, por llevarnos a los clandestinos más chungos de los Madriles, y haberse inventado la mejor canción a palmas. Y sobre todo gracias al Shooting Star que es **Luis**, por esta conexión tan guay, su buen rollo y su buena música, por más trambolikeo, y porque siga saliendo con su cara en las fotos; y a **Laura**, por su espíritu hiperluchador, su sentido de la justicia y la crítica, sus abrazos, y porque es un amor de persona (lo digo yo y lo dice Nacho). **Jesús, Dana, Andrea, María Gracia, María(s), Jorge**… no hay ni uno que no contribuya a que un training con vosotros equivalga a soltar más risas en una semana que el resto de meses del año juntos. Con vosotros me

he reído tanto como he aprendido. Y a **Eli**, por ser un amor de mamá de la promoción de becarios y haber sido otra más del grupo.

Existe un contexto ecológico para esta tesis, y ese contexto es Barcelona. Gracias a **Luis** y **Laura**, de nuevo, por esas salidas aquí y allá, esporádicas pero esenciales, por abrirme las puertas a sus casas. "Barça" (así la llamaba Floor) es una ciudad muy grande donde conoces a mucha gente, pero no es fácil dar con fauna autóctona. Gracias a **Mariona** (e indirectamente a la susodicha beca de la Caixa) por habernos juntado a raíz de un email y que surgiese la que es mi amistad más querida con alguien de Barcelona; gracias a **Mario** y **Alex** también, por muchas cenas y planes de sushi, Nintendo, charlas de arqueología y cotilleos varios hasta las tantas, y por el guerrero más letal. Gracias a **Cris**, por aparecer de la maldita nada en nuestras búsquedas de pisos y convertirse en alguien tan cercana; gracias a **Judith**, a **Flor**, a **Luis M**, y a otros compañeros de piso. Gracias a **Adriá**, por haber ayudado muchas veces a que no se apague la llama, y **Carles**, por todas esas quedadas hablando de ciencia (ficción). Thanks SO MUCH to **Aurélien** (Aurelio!), **Floor**, **Pili C** and **Khizer** (Jizer? Jisr? Whatever man), for having turned my first year in Barcelona into a spectacular experience. I had so much fun and was so fond of you guys! Thanks for the trip around your country **Aurélien**, and to your family too for being such nice hosts. Thanks for the crazy night in Lille with **Floor**, and for all the fun moments, good wine, fuet, cheese and quiche we had during that year! Long live Le Châtelet, the White Russians, and Long Live Alchemy! Gracias a **Matteo**, el compañero samurái durante un tiempo crucial en Barcelona. Grazie mile per tutti las muchísimas noches al balcón con el palet, la cerveza, las historias del tiburón, la música chill y los camboyanos. Gracias por la cocina, la pasta VERA, el fuet divino y el frutti di mare. Tendré las experiencias Voyager siempre en mi corazón. Ojalá cumplas tu sueño del hostel con lago de pesca… Buen viaje en las tierras del Hobbit!

Pero este viaje tiene raíces en aquellos años durante la carrera, e incluso en años atrás… Gracias a **Julia**, la jefita, y a **Eduardo**, por mis primeras lecciones en el trabajo de laboratorio, por escuchar mi entusiasmo por los protists y conducirlo a tardes de experimentos productivos. A **Julina**, por haber sido mi apoyo más importante en un momento tan crítico de mi vida, por haber hecho una aventura y una fiesta hasta de las tardes de limpieza, por El Pequeño Eucariotín, el Zorimushi, y por transmitirme el espíritu de perseveración. Jamás se me olvidará nada de eso. Al **Eduardo** por abrirme las

puertas de su lab, dejar que me agenciase pantallas viejas de ordenador, que colgase dibujos de paramecios por ahí, por haberme regalado separatas y fotocopias de hace décadas, y sobre todo por haber sido mi primera referencia de un investigador. Gracias al **departamento de micro**, y a **Abraham**, **Lin**, **Pedro**, **Mamen**, **Luis L**, **Pigne**, y a todos los demás, por todos los ánimos durante aquel año que fue crucial para decidirme por este camino y dar el salto a Barcelona. Fue, definitivamente la mejor decisión. Gracias a compañeros de residencia (**Ramiro**, **José Carlos**, **Jesús, Ángel, Antonio**, **Carlos**, **Jorge**) y a profesores que fueron una auténtica inspiración, y que vieron en mí el potencial de acabar siendo científico. **José María**, **Alfredo**, **Andrés**, **Blas**, **Rosa**, **Francisco Fernández**… lo que aprendí de vosotros significa mucho para mí.

Gracias a **Chema**, por reunirnos en capitales varias del viejo mundo a tener charlas cada vez más trascendentales; y al señor Wang, claro está. A **Pili**, por su amistad y su apoyo durante todos los años de la carrera. Gracias a **Kiko**, que ha estado ahí literalmente desde siempre y jamás dudó de mí y de decirme las cosas claras cuando hacía falta; y a **Iván**, por la "prueba de tener un corazón de acero" y otras mil historias que él sabe tan bien.

Y gracias a mi familia.

A mi **familia de Badalona** (**Pepe**, **Maricarmen**, **Michel**, **Lola**, **Ani**, **Pepe**, y todos los primos), por haberme acogido al llegar. A mis tíos: a **Chari**, **Ani**, **Cristina**, **Mariche**, y a las primas, por haberme apoyado desde la distancia. A mi tío **Arturo**, por, junto con mi padre y mi tío **Manolo**, haberme inculcado el amor al monte y a la naturaleza. Por todas esas excursiones de toda la familia juntos. A mis **abuelos**, por las bromas del espíritu **Pérez**, y por el carácter y lo asertivo del **Posada**, porque ojalá sientan que este nieto ha llegado adonde se esperaban. Les mando el beso más grande del mundo, a los que están aquí y a los que ya están allí.

A mis hermanos **Miguel** y a **Carlos**, por compartir el gusto por los bichos raros, las mismas bromas, las mismas tonterías, porque me apoyáis sin que os deis cuenta, porque cada vez que nos vemos es como si no hubiera pasado el tiempo. Pero sí que pasa, estáis enormes, y más grandes os quiero ver aún.

Gracias a mis padres: A ti, **papá**, por haber respondido siempre a todas mis preguntas de animales hasta que un día empecé yo a responder las tuyas, por los acuarios, las iguanas, los tritones y el Monte de la Rana; a ti, **mamá**, por haberme inculcado el trabajo constante, la dedicación y el esfuerzo, y haber sabido escuchar y aportar consejos en todos los momentos de mi vida. Ambos me habéis apoyado y recogido cuando he caído, así que gracias por todo ello.

A **Nacho**. A mi hermanito. Haber llegado hasta aquí es un orgullo para mí mismo, pero me agrada saber que para ti también. Gracias por todo en este tiempo. Por coger muestras conmigo para mirar al microscopio, por estudiar juntos las Hydras y los rotíferos en bachiller, por ahogarnos de la risa con *“En qué Idioma lloran los Caracoles”*, por las conversaciones hasta las tantas durante más de diez años, por alimentarnos mutuamente esa llama de curiosidad. Por haber sido de los primeros en haber bromeado con que me acabaría convirtiendo en un “Doctor en biología de microorganismos” allá por primero de carrera (aunque él quizá no se acuerda, pero no importa, porque ya me acuerdo yo). Por medio saber qué pensamos con tal de decir dos cosas, por más que pasen los años. Great minds think alike, but so do good people. A fin de cuentas, he concluido que por eso “nos parecemos tanto”. *ECAEMYDECC*. [illegible].

Y a **Carmen**. Por habernos reencontrado en un momento crucial, por haberle dado la vuelta a todos los esquemas de mi vida, y haberme ayudado desde el primer minuto a reencontrarme a mí mismo y haberme dado paz. Te debo muchísimo por todo este tramo final. Menos mal que cuento con la vida entera para ir devolviéndotelo. Road work ahead? I sure _know_ it does! *Sarangmyeon*!

Abstract

The origin of animal multicellularity has its roots in the process by which one cell becomes many: the cell division, in turn controlled by the cell cycle. Many features of animals, such as development, rely in an intricate control of the cell cycle. The study of the evolution of animal cell cycle has the potential to shed light into the foundations of this lineage; however, cell cycle regulation has been characterized in few species belonging to other evolutionary lineages, for which comparative approaches are limited. In this thesis I present the work leading to the development of tools for cell cycle synchronization and DNA transfection in the species *Capsaspora owczarzaki*, a unicellular relative of animals proposed as a model system to understand the transition to metazoan multicellularity. These tools have allowed to infer some key aspects of the cell cycle regulation in the unicellular ancestor of animals, and open the door for a new era of functional approaches to address the question of the origin of animals.

Sinopsis

El origen de la multicelularidad en animales tiene sus raíces en el proceso por el cual una célula da lugar a otras: la división celular, controlada a su vez por el ciclo celular. Muchas características de los animales, como el desarrollo embrionario, requieren de un control preciso del ciclo celular. El estudio de cómo evolucionó la regulación de ciclo celular en animales tiene el potencial de arrojar luz a los orígenes de este linaje; no obstante, la regulación del ciclo celular ha sido caracterizada en pocas especies pertenecientes a otros linajes evolutivos, lo cual dificulta una aproximación comparativa. En esta tesis, presento el trabajo que ha dado lugar al desarrollo de herramientas para la sincronización de ciclo celular y la transfección de ADN en la especie *Capsaspora owczarzaki*, un pariente unicelular de los animales propuesto como organismo modelo para entender la transición a la multicelularidad. Estas herramientas han tanto permitido inferir aspectos esenciales de la regulación del ciclo celular en el ancestro unicelular de los animales, como abrir las puertas para una nueva era de aproximaciones funcionales para abordar la pregunta del origen de los animales.

Preface

As many other biologists before me, yours truly hasn't been the first (nor will be the last) to try and explain to himself what 'life' is (or what is it about), and would struggle to find a way to describe it in simple terms. In retrospective, the closest I have come to this is a simple list of three tightly interrelated concepts: diverse, responsive, and self-perpetuating. 'Diverse' because, despite the common building blocks that are biomolecules and cells, we observe a huge diversity of extant life forms: microscopic bacteria and unicellular eukaryotes, photosynthetic plants and algae transforming the energy of the Sun and making this planet hospitable, animals of all kinds predating on other organisms and among themselves, all of them occupying all niches and places one can think of as inhabitable (and even those ones we wouldn't). Living organisms are responsive as dynamic systems capable of perceive changes in an interacting environment and accordingly react at multiple levels (from physiological to behavioural). Along the same line, species are not static because they are susceptible to changes and alterations that confer them greater fitness over successive generations in time. This occurs as living forms self-perpetuate through reproduction, either at the unicellular or multicellular levels, either asexually or sexually. All these traits are linked to each other by another pivotal concept, the genetic program, in an interplay with consequences at the evolutionary time scale.

It is nothing but mesmerizing that the grandiose scenario that is life on Earth can be reduced to a minimal unit of function, the cell. Because all known life descends from a single ancestral cell, all life originates from a single-celled state, and a notorious variety in all life is explained by differences in the arrangement and types of cells; all of it underlied by the ever-running mechanism of cell reproduction, in turn regulated by a genome at the mercy of evolution's whims.

Let us imagine, for a moment, a gene within a cell. That is, a sequence of DNA; a molecule formed by nucleotides, that is, smaller components. May this gene's product be linked to a certain feature. May this feature be some sort of chemical activity within the cell, may it be the ability to respond to a cue, may it even be the ability to give rise to a cell lineage, responsible to give rise to an essential component within a multicellular body plan.

Let us think, for a moment, about the many generations and processes of speciation during so many millions (I say again, 'millions') of years so that, just by chance and only chance, some changes in this molecule occurred, that some -but not all- of the changes were translated into a certain other feature, at first unnoticed, first in one organism, then transmitted to some, then to many, to the point of being present in all and thus becoming subjected to selection; that in turn this sequence gathered more changes, that some -but not all- provided a larger fitness, and were thus transmitted first to some, then to many, to the point that it became a normal feature in the species. Only to start over again. Let us think, for a moment, about this same process, only for all the features, all those details at the structural, metabolic, morphological and behavioral levels, that make a species being what it is.

Hundreds, thousands, hundreds of thousands, perhaps millions of changes in the tiniest molecules, inappreciable molecules, for millions of years. And yet, as small as they are, they are so important. They determine so many things. They made you to be here, alive, sitting and about to read -and hopefully enjoy- my thesis, for example. And not only us as a species. Let us think, for a moment, in all those changes but only for all the species, all the known and unknown species. All the organisms that were born, that were no more, that may have thrived, that may have not.

As many other biologists before me, yours truly is one who fascinates when drifting away in this sea of thoughts, because when one stops and thinks in all the 'genes within a cell', in all the cells, in all of Earth's history, it is inevitable to end up admiring the complexity and might of life, no matter the definition. To attempt and abstract the immensity of evolution of life on Earth down the level of cells, genes, and molecules, is a precious exercise to a researcher. An exercise that makes you feel very, very small.

But, at the same time, happy and proud to be part of such a big thing.

This thesis is my humble attempt to sail this sea of thoughts. Now, if you sail with me…

Prefacio

Como muchos otros antes que yo, este biólogo al que lees no ha sido el primero (ni será el último) en intentar darse a sí mismo una respuesta de qué es 'la vida' (o de qué se trata), y dudaría en cómo sería capaz de describirla con palabras sencillas. En retrospectiva, lo más que me he acercado a esto es una lista de tres conceptos muy relacionados entre sí: diversa, responsiva, y que se perpetúa a sí misma. Diversa porque, a pesar de compartir los ladrillos básicos que son las células, los genes y las moléculas, vemos una diversidad enorme de formas de vida: bacterias microscópicas y eucariotas unicelulares, algas y plantas fotosintéticas transformando la energía del sol y haciendo de nuestro planeta un lugar acogedor, animales de todas clases depredando a otros organismos, todo ello ocupando todos los nichos y lugares que uno pueda imaginar como habitables (y probablemente también aquellos impensables). Los seres vivos son responsivos en tanto que son sistemas dinámicos capaces de percibir cambios en un ambiente interactivo, y reaccionan acordemente a múltiples niveles (desde fisiológicos hasta de comportamiento). En la misma línea, las especies no son estáticas ya que son susceptibles a cambios y alteraciones que les confieren mayor aptitud conforme las generaciones se suceden unas a otras. Esto ocurre al tiempo que los seres vivos se auto perpetúan al reproducirse, ya sea a nivel unicelular o multicelular, ya sea a nivel asexual o sexual. Todos estos atributos están vinculados entre sí gracias a otro concepto igual de importante, el programa genético, en un sistema de interacciones con consecuencias a escalas evolutivas.

Es inevitable verse cautivado por el hecho de que, tan grande como es el escenario de la vida en la Tierra, pueda reducirse a una unidad mínimamente funcional, la célula. Porque toda la vida conocida desciende de una célula ancestral, toda la vida comienza como un estadio unicelular, y gran parte de la diversidad de la vida se puede explicar por diferencias en la organización de las células. Todo ello subyacido por el incesante mecanismo de la reproducción celular, a su vez regulado por un genoma a merced de los caprichos de la evolución.

Imaginemos, por un momento, un gen perteneciente a una célula. Una secuencia de ADN, una molécula formada por nucleótidos, es decir, componentes más pequeños. El producto de este gen está asociado a cierta característica. Esta característica puede ser desde una actividad química en la célula, a la habilidad de responder a cierto estímulo,

hasta la habilidad de dar a una progenie de células que formarán un componente esencial en el cuerpo de un organismo multicelular.

Pensemos, por un momento, la de generaciones y especiaciones que habrá habido durante tantos millones (repito, millones) de años para que, por azar y sólo azar, hubiera cambios en esas moléculas, que algunos (no todos) se tradujesen en cierta característica, inicialmente inapreciable, en primero uno, luego varios, luego muchos organismos, hasta el punto de estar siempre presente y por tanto sujeto a selección, para que a su vez aunase más cambios, que resultasen beneficiosos y se transmitiesen a tal cantidad de descendientes de la población que pasó a ser una característica habitual en la especie. Y vuelta a empezar. Pensemos, por un momento, en eso mismo, pero para todas aquellas características, todos esos detalles, a nivel morfológico, estructural, metabólico, … que hace a cierta especie ser la que es.

Cientos, miles, cientos de miles, quizá millones de cambios en moléculas chiquitísimas, inapreciables, durante millones de años. Y sin embargo, por muy inapreciables que sean, son tan importantes. Determinan tantas cosas. Han hecho que estés aquí vivo, sentado y leyendo mi tesis, por ejemplo. Y no sólo a nosotros como especie. Pensemos en esa cantidad de cambios, pero para todas las especies conocidas y por conocer. La de organismos que habrán nacido, que habrán muerto, que habrán proliferado, que habrán perpetuado sus genes, que no lo hayan conseguido.

Como muchos otros antes que yo, este biólogo al que lees es de los que se fascina al divagar en este mar de pensamientos, porque cuando uno se para a pensar en todos los ejemplos de situaciones como la de hoy es inevitable acabar admirando la complejidad y poder que tiene la vida, la definas como la quieras definir. Intentar abstraer la inmensidad de la evolución en la Tierra pormenorizando en cada detalle hasta llegar al nivel de células, genes y moléculas es algo precioso. Y te hace sentir muy, muy pequeño.

Y al mismo tiempo orgulloso de formar parte de algo tan grande.

Esta tesis es mi humilde intento de navegar ese mar de pensamientos. Y ahora, si te animas a navegar conmigo…

Table of contents

1. Introduction

Introduction

Evolutionary biology focuses on understanding the events and processes that led to the current biodiversity and complexity of living organisms. Life as we know it has always been subjected to change by evolutionary forces, but the most notorious radiations have occurred around the so-called major evolutionary transitions (Rokas, 2008; Szathmary and Smith, 1995). These milestones led from simpler self-replicating entities into higher ones: the origin and confinement of replicating molecules, and the endosymbiosis of prokaryotes giving rise to a new cell architecture, are only some of them. One, in particular, stands out due to its multiple appearances across very distant lineages: the transition from unicellular to multicellular levels of organization in eukaryotes. In multicellular eukaryotes, single cells assemble to form an integrated, coherent collective participating as a whole entity in the evolutionary process. Multicellularity per se comes hand in hand with emerging properties allowing to create and occupy new ecological niches in ecosystems, resulting in an enrichment of the biodiversity. The mechanisms by which these emerging properties can be acquired and are regulated are only vaguely understood, a reason for which is necessary to define a framework of the rationale underlying multicellular systems.

1.1 Overview of the evolution of multicellularity

Multicellularity has appeared in a plethora of different lineages, with remarkable mentions to the embryogenic multicellularity of land plants, brown algae, fungi, and animals (Adl et al., 2019; Grosberg and Strathmann, 2007; King, 2004; Knoll, 2011) (Figure 1A). Albeit these lineages are very different among them (with their common unicellular ancestor dating back to the radiation of the major groups of eukaryotes, (Adl et al., 2019)), they share numerous common aspects. First, they (especially land plants and animals, Fig. 1D) account for representing a large amount of the diversity in morphologies and lifestyles we observe (Bar-On et al., 2018; Mora et al., 2011). Despite all these differences, all of them are ultimately formed by cells that originate from one zygotic cell, meaning all the cells in a multicellular organism contain (in principle) the same genetic

information and can arrange and coordinate to form a coherent body plan following a developmental program (Figure 1C). These arrangements culminate in systems and tissues that cooperate to maintain the homeostasis of the whole, with specialized cell types simultaneously coexisting within the same organism. Unicellular organisms exhibit a contrasting ontogeny: they emerge from single cells but do not follow a developmental plan. Instead, they temporally transition between life stages widely depending on external cues (such as the availability of nutrients or other triggering stimuli), because the co-occurrence of different biological processes within the cell is not always possible (Figure 1B, 1C). Additionally, unicellular organisms need to cope with the colliding two essential processes that are gene expression and DNA replication; in other words, they need to integrate the responses to stimuli and the execution of gene expression programs while they proliferate and replicate their DNA (Pomerantz and O'Donnell, 2010). These constraints put unicellular organisms under strong selective pressures for optimal response to the surrounding context.

Multicellular organisms overcome these limitations typical in unicellular organisms: spatial differentiation of cell types allows different tasks otherwise incompatible to occur at the same time in a single entity. For example, different metabolic pathways like nitrogen fixation and photosynthesis are separated in two different cell types in cyanobacteria (Fay and Walsby, 1966), machinery-competing processes like mitosis and motility can occur in flagellate colonies as some cells swim while others generate more cells (Buss, 1988), and the aforementioned replication-transcription colliding is overcome by cells with heavily-specialized programs performing unique tasks while delegating the process of DNA replication and mitosis to other cell types. Multicellularity also comes with the advantages of a larger size: organisms can escape predators but also enhance their predatory capabilities, and in the case of autotrophs it provides larger lit surfaces (Boraas et al., 1998; Pentz et al., 2015; Ratcliff et al., 2013) (Figure 1C).

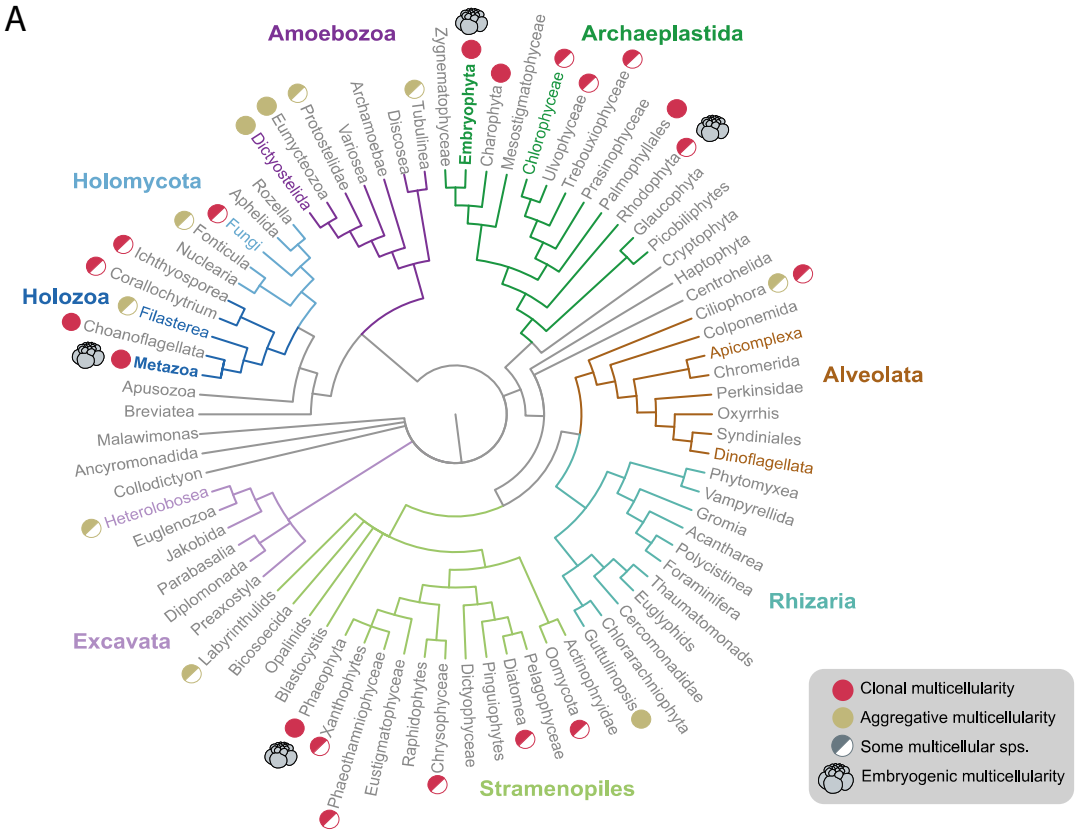
Together, all the mentioned arguments are seen as possible adaptive scenarios where multicellularity can develop. In itself, they cannot explain why the transition occurred in the first place for two reasons: first, they are rather Darwinian properties and therefore cannot be invoked to explain their own evolution (Clarke, 2014; Griesemer, 2001; Libby and Rainey, 2011; Rainey, 2007; Rainey and Kerr, 2010); second, selection at the lesser level of organization (cells) can disrupt the selection process at the higher level during the transition to multicellularity (De Monte and Rainey, 2014; Szathmary and Smith, 1995) (Figure 1B). A mid-level of selection takes place when all cells within a collective

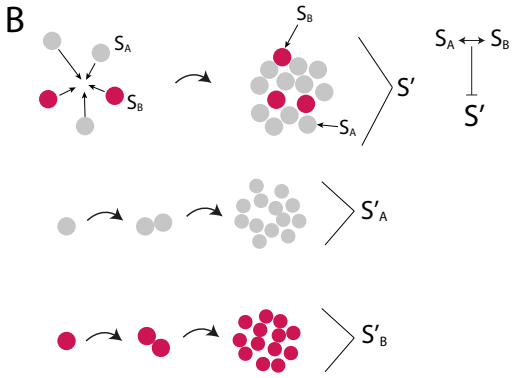
come from the same parental cell: they all have the same fitness level, but fitness across collectives is different. This results in a switch from within-group to between-groups selection (Figure 1C). Only then the groups become evolutionary units and obey natural selection (a detailed description is available in (Szathmary and Smith, 1995) and (De Monte and Rainey, 2014)). At some point the forming cells lose the ability to replicate on their own and require the context of their higher level of organization (as in specialized cell types depending on proliferative signals in tissues; see (Szathmary and Smith, 1995) for more examples), due to reasons often not related to the evolution of such complexity in the first place. This happens only in organisms with a pre-adaptive value in their regulatory program, that is, the genome content. In turn, this pre-adaptive value comes from a byproduct of unidirectional changes (mutations) in regulatory systems where their parts, originally non-dependant, become increasingly co-dependant, in a two-scale scenario of contingent irreversibility (Gray et al., 2010; Lukeš et al., 2011).

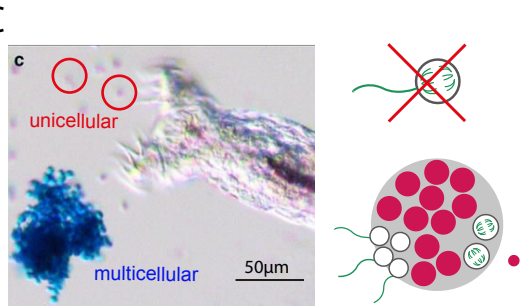
We can think of the origin of animals as a large-scale, illustrative example of pre-adaptiveness and the acquisition of complexity in the transition to multicellularity. The appearance of new genes and regulatory features involved in different aspects of metazoans (such as body plan and development) has been linked to the origin of animals, including Wnt, TGF-Beta, animal-like enhancers, different histone marks, the presence of different networks and interactions, or even the transition between cell types. However, numerous studies in comparative genomics have shown that the unicellular ancestor of animals was already complex (meaning many of these genes and regulatory features were already present). The transition to (and irreversibility of) animal multicellularity occurred not because of selection pushing to acquire more complexity, but simply because many of the components later part of the "multicellular toolkit" had been selected for other reasons, their interactions ambiguous or non-deleterious to the eyes of natural selection.

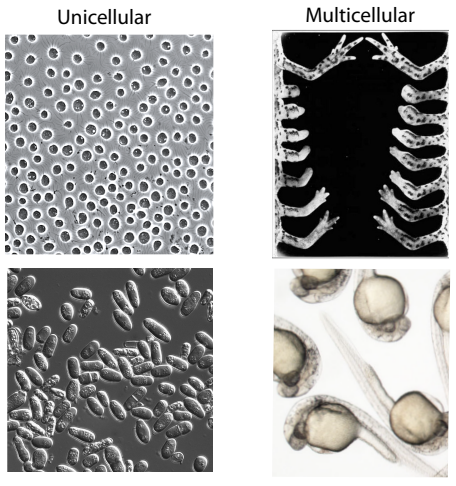
This “two-level” (genomic and cellular), contingent irreversibility allows to explain why multicellularity has appeared more than once (their unicellular ancestors with a certain degree of complexity in their systems) but neither always nor the same way (other lineages are under strong selective pressures, their regulatory mechanisms constricted and with a limited margin for evolution). Parallel studies in fungi and plants complement this perspective: the expansion of transcription factors in Embryophyta took place for different families than animals (de Mendoza et al., 2013), and the formation of fungal multicellular structures in Dikarya likely did not involve attachment of daughter cells but rather the innovation of septation mechanisms in long, multinucleated hyphae (Nagy et al., 2018). In general terms, the acquisition of multicellularity is related to the ability to keep cells together, of cells to communicate between them, and of cells to acquire specialized functions. The current knowledge suggests that this can happen at different levels of regulation by using very different molecular toolkits, including rewiring of cell division regulation, expansion of pre-existing signaling genes and transcription factors, and the transition from temporal-only to temporal-and-spatial cell differentiation (Mikhailov et al., 2009; Szathmary and Smith, 1995).

Figure 1 (next page). An overview of the origin of multicellularity. A: Phylogenetic tree of eukaryotes showing the different lifestyles of multicellularity (adapted from Sebé-Pedrós et al., 2017). Colored tree leaves indicate the main lineages where the cell cycle has been studied and/or characterized. B: The transition to multicellularity: Complex multicellularity such as those of animals and landplants can only be achieved when multicellular structures become evolutionary units. This can occur only if cells within a group have all the same fitness (S'_A, S'_B), meaning there is no within-group selection that could disrupt selection processes between collectives (S_A, S_B, S'). Red and gray circles represent cell with different genetic information and therefore different fitness. C: multicellular organisms (shown in blue) can escape predators (such as the rotiferan in the picture) more efficiently than unicellular organisms (single cells stained in red), and can couple tasks otherwise incompatible (such as mitosis and flagellar motility). Adapted from (Pentz et al., 2015). D: Unicellular organisms often do not follow a complex ontogenetic cycle, instead transitioning between different stages and reproducing in response to environmental cues. Cells within multicellular organisms divide and proliferate under certain conditions according to the needs of the whole organism, such as tissue or organ regeneration, and embryonic development, by integrating various signals from the multicellular environment in which they are found. Clockwise: *Capsaspora owczar*zaki adherent cells adapted from (Suga et al., 2013), Salamander limb regeneration (Goss, 1969), Zebrafish embryo, and *Schizosaccharomyces pombe* cells (Wikimedia commons). E: different representatives of multicellular eukaryotes across the tree of life. Clockwise: a mushroom (Fungi) and mosses (Embryophyta) (Wikimedia commons), a colonial individual of *Volvox* (Wikimedia commons), the sea anemone *Nematostella vectensis* (Putnam et al., 2007), and the amoebozoan *Dictyostelium discoideum*.


A
Amoebozoa
Archaeplastida
Holomycota
Holozoa
Alveolata
Rhizaria
Excavata
Stramenopiles
Fungi
Metazoa
Embryophyta
Chlorophyceae
Dictyostelida
Heterolobosea
Dinoflagellata
Clonal multicellularity
Aggregative multicellularity
Some multicellular sps.
Embryogenic multicellularity


B
S′
S′A
S′B


C
unicellular
multicellular
50µm


D
Unicellular
Multicellular


E

1.2 Cell division underlies multicellular processes

As pointed by (Szathmáry, 2015), the phenomenon underlying the aspects of multicellular life previously discussed (and the same necessary to acquire them) is none other than the control of cell division. Something inherent to life's self-perpetuation is the need to preserve the information as the organisms reproduce, in turn meaning that faithful replication and splitting of the genetic material must be ensured when cells divide. Tight regulation of cell division and proliferation is what drives populations of microorganisms to proliferate in optimal environmental conditions, to transition between life stages, to establish cell types generating specific responses, or the formation of intricate biological structures like body plans in an interplay with the developmental program. Cell division is also the mechanism to either directly generate offspring or to generate the cells that will give rise to new organisms, implying it is the biological window in where changes in the genetic information can be fixed and preserved over time, with potential evolutionary implications. Differences in the way cell division operates between species (including unicellular and multicellular) are explained by differences in the control of the cell division process through the genetic regulatory program. Lastly, recent evidences (discussed below) show that tweaking the genetic program of cell division can have drastic consequences in the emergence of simple, conglomerate multicellularity.

A better comprehension of how cell division is controlled by the genetic program across different taxa can ultimately impact our knowledge on how the transition to multicellularity can happen. For this reason, we need to investigate the process responsible for regulating the proper consecution of the different events required for cells to divide: the cell cycle.

1.3 The cell cycle

a) An essential order

DNA replication and mitosis ensures proper cell reproduction

In the nineteenth century, a contemporary of Charles Darwin named Rudolph Virchow introduced the third and then final dictum to the cell theory: *Omnis cellula e cellula*; all cells come from cells (Kuiper, 2010). As previously discussed, all living things (including multicellular ones) come from pre-existing cells and descend from a single unicellular ancestor. The only way to make a cell is by duplicating a previously existing cell, meaning all the material and information in the mother cell needs to be correctly copied and transmitted to the daughter cell. This is ensured by timing the duplication and splitting of the genetic material. The resulting rounds of replication and division are known as the cell cycle, and comprise the essential underlying process of reproduction.

Controlling the elaborate process of replicating the genome and dividing the cell is facilitated by temporally spacing such events into different cell cycle phases of DNA Synthesis (S phase) where chromosomes are duplicated, and Mitosis (M phase) where they are segregated. In many cases, these events are separated by two gap phases, namely G1 (between M and S) and G2 (between S and M), where the cell synthesizes different components, multiplies their protein and organelle content and grows in size (Morgan, 2007). These phases also offer the cell a time window to inspect their internal and external environment for suitable conditions, such as the completion of all the previous cell cycle events or environmental cues, before committing to the next one (Figure 2 A).

G1, S, and G2 are together called interphase, as opposed to mitosis. After division, a newborn cell usually grows larger and undergoes active metabolism (read this and make sure) during G1. On a certain moment in this phase, the cell can decide if continuing proliferation or exiting the cell cycle. If the latter was the case, cells would enter into a so-called G0 state or no-return point where they undergo a reprogramming of gene expression, for example giving rise to specialized cell types. Sometimes these cells never re-enter the cell cycle and stay in G0 until they die. If the cell decided to follow further rounds of duplication, several mechanisms are activated to increase cell size and to start preparing the cell for DNA replication. This moment is called Start in yeasts, and

Restriction Point in mammalian cells. During S phase, each molecule of DNA, or chromosome, is replicated and chromatin proteins are synthesized to provide a scaffold for the nascent DNA. Centrioles, a particular organelle with a significant role in the upcoming stages, are also duplicated in S phase. After DNA replication is completed, the cell is committed to continuing into G2 and mitosis. Microtubules assemble around the centrioles forming centrosomes, and organelles and the membrane system dissociate, including the nuclear membrane, and scatter across the cytosol.

Upon entering cell division, newly-replicated chromosomes are condensed in densely-packed structures called chromatids and kept together by cohesins in prophase. A macromolecular complex called kinetochore forms on the centromere, the primary constriction of chromosomes. The emanation of microtubules from the centrosomes originates the mitotic spindle, whose filament tip proteins associate to the kinetochores of sister chromatids. The mitotic spindle thus positions centrosomes and the chromosomes relative to the plane of division forming the metaphase plate, facilitating the subsequent segregation of chromosomes (Figure 2A-B). The tensions of the mitotic spindle to pull apart chromatids are compensated by the proteins keeping sister chromatids together. In anaphase, a protein signaling cascade triggers the dissociation of sister chromatids, resulting in spindle filaments pulling one chromatid of each chromosome to either pole of the dividing cell. These structures finish resolving in telophase, where the spindle dissociates, and the different organelles are restituted. These events comprising nuclear division, also called karyokinesis, are followed by division of the cytoplasm or cytokinesis. With variances across cells and species, an actomyosin complex forms in anaphase beneath the plasma membrane of the equatorial plane, pulling the membrane inside by gradual contraction. Cytokinesis concludes by fusion of plasma membrane sealing the gap between the daughter cells.

Each of the major events in this sequence is binary in response because of their implications: once started, the cell is committed to finish. An example is found in cells that become refractory to extracellular factors once they have entered cell cycle (Hartwell et al., 1974). The disassembly of the nuclear membrane and the condensation of chromosomes would disrupt homeostasis if left uncompleted. Similarly, uneven segregation of chromosomes due to incomplete DNA replication would yield a newborn cell lacking part of the genome, with catastrophic consequences. Concordantly, proper cell volume needs to be ensured, meaning checking all the organelles and other cellular material are abundant enough to sustain the newborn cells, before starting DNA

replication and mitosis. The essential order and pacing of cell cycle phases contributes thus to a tight regulation of the cell cycle.

Because cells can appear in a multitude of contexts, the cell cycle is deployed differently on every situation. Two examples are palintomic divisions in vertebrate embryonic cell cycles lacking gap phases, or the modified cell cycle in meiosis where two rounds of division follow one round of duplication. These adaptations are ultimately variants at the robust, multi-layered regulation of the cell cycle, whose aspects are discussed below.

b) The core control System: CDK-cyclin complexes

The fundamental regulatory machinery of the cell cycle is a biochemical oscillator of protein phosphorylation, ubiquitination, and degradation activities (Morgan, 2007). The central components of this oscillator are a group of protein kinases deeply conserved in eukaryotes named cyclin-dependent kinases (Lee and Nurse, 1987; Reed, 1980). These kinases control the activity of many other proteins involved in cell cycle control, including the activation or inactivation of the main effectors of the cell cycle (Dunphy and Newport, 1988; Fisher et al., 2012; Hunt, 1989; Lohka, 1989; Peter et al., 1990). As kinases, they have the typical two-lobed conformation, with different modifications for further regulation (Fig. 3). A flexible portion of their structure, named T-loop, regulates the binding to protein substrates (Fig. 3). Another important region is one composed by the two alpha helices containing a “PSTAIRE” motif, which interacts with the main partners of CDKs regulating their activity over the cell cycle (Fig. 3) (Brown et al., 2015; De Bondt et al., 1993).

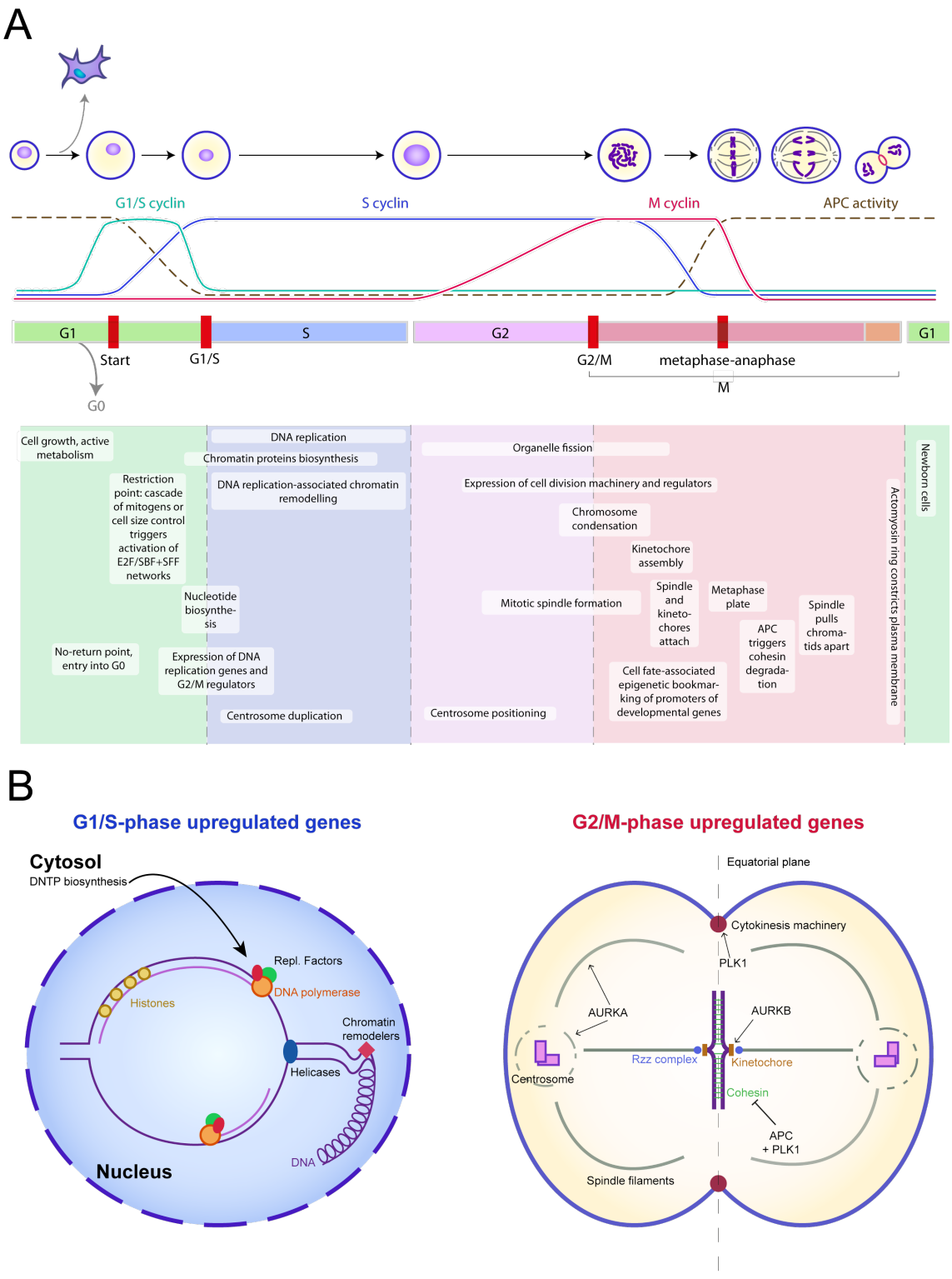


Figure 2 (next page): The eukaryotic cell cycle. A: 'bird's-eye view' approximation to the eukaryotic cell cycle. White boxes indicate the different events occurring approximately at each time, in an approximate way to all analyzed model species. B: an overview of some of the important genes regulating two key steps of the cell cycle: DNA replication and cell division.

The regulation of the cell cycle, therefore, consists of gradual increases and decreases of activity mediated by binding of CDKs to another group of regulatory proteins called cyclins (Peeper et al., 1993). Protein levels of cyclins oscillate over the cell cycle by transcriptional regulation and protein degradation (Botz et al., 1996; Cross et al., 1999; Evans et al., 1983; Kobayashi et al., 2015b; Le Cam, 1999; Ohtani et al., 1995; Polanowska et al., 2001), constituting a great example of biological timers in nature. Cyclins are not that much conserved in sequence, but they all share a tertiary structure of two compact domains with five alpha helices each. The first compact domain corresponds to the cyclin box, a region of 100 amino acids where most of the sequence similarity between cyclins concentrates. Sequences outside this box are variable, especially the amino-terminal region containing different regulatory and targeting domains specific for each cyclin (Brown et al., 1995). Because this fold is also present in other transcription factors such as retinoblastoma or TFIIB, it is possible that cyclins share a common evolutionary origin with some proteins mediating transcription (Gibson et al., 1994).

Binding of cyclins to CDKs modify the protein conformation of the CDK active site, increasing the binding affinity of CDKs for their substrate (Peeper et al., 1993). Each cyclin also provides a different degree of affinity to the substrate, which implies that different cyclins activate different downstream targets. The range of cyclin-CDK targets is expanded by indirect binding of CDKs to substrates via cyclin interactions, or by regionalizing CDK activity to a subcellular compartment by co-localization with cyclins.

Furthermore, cyclin-CDK activity control extends beyond the cell cycle. In animals, complexes of both non-canonical and canonical cell cycle cyclins and CDKs have specific tissue expression and have been reported to play roles in RNA polymerase II- and Wnt/β catenin- based transcription, DNA damage repair, metabolic and epigenetic control, or establishing cell fates (Lim and Kaldis, 2013). In yeasts, Cln3 is a cell size controller with no periodic activity (Richardson et al., 1989; Wang et al., 2009).

Because each complex of cyclins and CDKs are active at different stages of the cell cycle, they are named G1/S, S, or M cyclins and CDKs, although this classification does not always apply. Under certain circumstances, such as different cell types, the same cyclin can participate in different cell cycle events. In some animal embryonic cells where there is no restriction point, G1/S cyclin levels remain high during the cell cycle and behaves as the S phase cyclin, while the otherwise S cyclin acts during M phase

(Morgan, 2007). Together with their many downstream targets, cyclin-CDK complexes of one phase also mediate the activation of complexes of the next phase, generating a robust regulation of progression of the cell cycle.

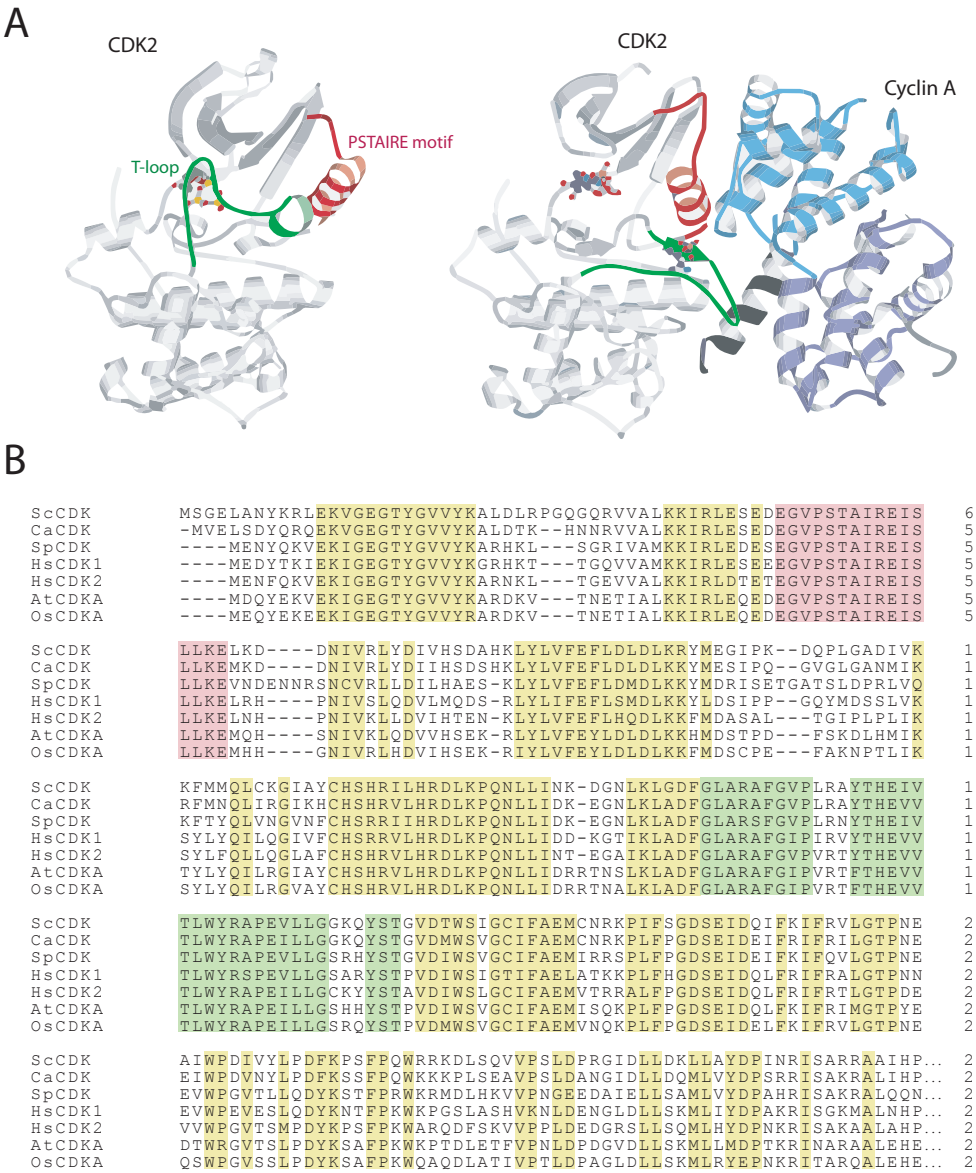


Fig. 3: Cyclin-dependent kinases, the drivers of the cell cycle. A: the three-dimensional structure of CDK2, an S-phase CDK in human cells. Red indicates the PSTAIRE motif, and green indicated the T-loop motif. Key residues in the active site are highlighted. When CDK2 interacts with cyclin A (far right), it suffers a conformational change and the active becomes accessible. Further activity is enhanced by CDK-activating kinases (CAKs). B: an alignment of seven CDKs sequences (three yeasts, two plants and two human orthologs). Highly conserved sites are highlighted in yellow. PSTAIRE and T-loop follows the same color scheme as A. Adapted from Morgan et al., 2004.

c) The transcriptional program during the cell cycle: Transcription factor networks and effector machineries

Both DNA replication and mitosis require a highly specific set of proteins and molecules that need to be present and expressed for such a ready-state of the machinery. Cell division cycles are linked to a temporal transcriptional program that ensures the presence of proteins when they are required. Different networks of transcription factors allow the "just-in-time" expression of many necessary downstream targets such as DNA replication or cell division proteins (Bendris et al., 2015; Doenecke, 2016). For instance, the E2 Factors - Retinoblastoma (E2F-Rb) network of transcription factors controls initiation of DNA replication in many eukaryotes (including animals) at the G1/S transition (Goodrich et al., 1991; Grant et al., 2013; Ishida et al., 2001) and are indirectly involved in the transcriptional regulation of G2/M by controlling the expression of other transcription factors, such as FoxM1 or B-Myb (Costa, 2005; Grant et al., 2013; Laoukili et al., 2005; Musa et al., 2017; Sadasivam et al., 2012). In yeasts, G1/S is driven by the SBF-MBF, Fkh1-2 and Hcm1 networks (Gefeng et al., 2000; Pramila et al., 2006; Reynolds et al., 2003). To a certain extent, these networks can run and self-regulate independently in parallel to progression through the cell cycle (Orlando et al., 2008).

In this thesis, the term "periodic transcriptional program" of a cell or a species refers to the catalogue of genes with oscillatory expression during the cell cycle of that cell type or species. The periodic transcriptional program typically groups in clusters of coexpressed genes that correspond to temporal waves of transcriptional activation, occasionally regulated by the same transcription factor (Fig. 4, Fig. 2) (Pramila et al., 2006). The two, most typical temporal waves found across species occur at the onset of DNA replication and mitosis, and are thus named the G1/S wave and the G2/M wave. The G1/S wave typically includes genes such as DNA polymerases, helicases, ribonucleotide synthases, histones, chromatin protein biosynthesis, and other proteins necessary to start and proceed through DNA replication (Whitfield et al., 2002) (Figure 2B). This temporal wave is mainly regulated by the E2F-retinoblastoma (Rb) complex (Fig. 4). The G2/M wave comprises mitotic spindle proteins, chromatin remodelers, and different regulators with a role in cell division: cohesins, keeping sister chromatids together; condensins, sister proteins of cohesins implicated in chromosome condensation; aurora kinase A, a serin-threonin kinase implicated in the function of the centrosomes and the mitotic spindle, aurora kinase B, a related protein involved in the

alignment and segregation of chromosomes; and polo-like kinase 1, a protein that indirectly triggers the separation of sister chromatids in collaboration to the Anaphase-promoting complex (see below) (Golan et al., 2002; Hansen et al., 2004; Kotani et al., 1998; Moshe et al., 2004) (Figure 2B). This wave is usually regulated by transcription factors from the Forkhead family, such as FoxM1 in animals or Fkh in yeasts (Fig. 4). Other biological processes that depend on cell cycle can also be subjected to periodic transcriptional regulation, such as metabolism, growth, synthesis of organelle components, development, cell differentiation, or programmed cell death (Cho et al., 2001; Kalucka et al., 2015).

Transcriptional activity during the cell cycle is present in numerous species and cell types, although many of the genes who are periodically expressed are unique to each. For example, it is common to find G2/M-upregulated genes related cytokinesis in every species, although the underlying mechanisms of cell division are different: mitosis and cytokinesis in animal cells, budding or fission in bud and fission yeasts, or cell wall deposition in *Arabidopsis* (Guertin et al., 2002).

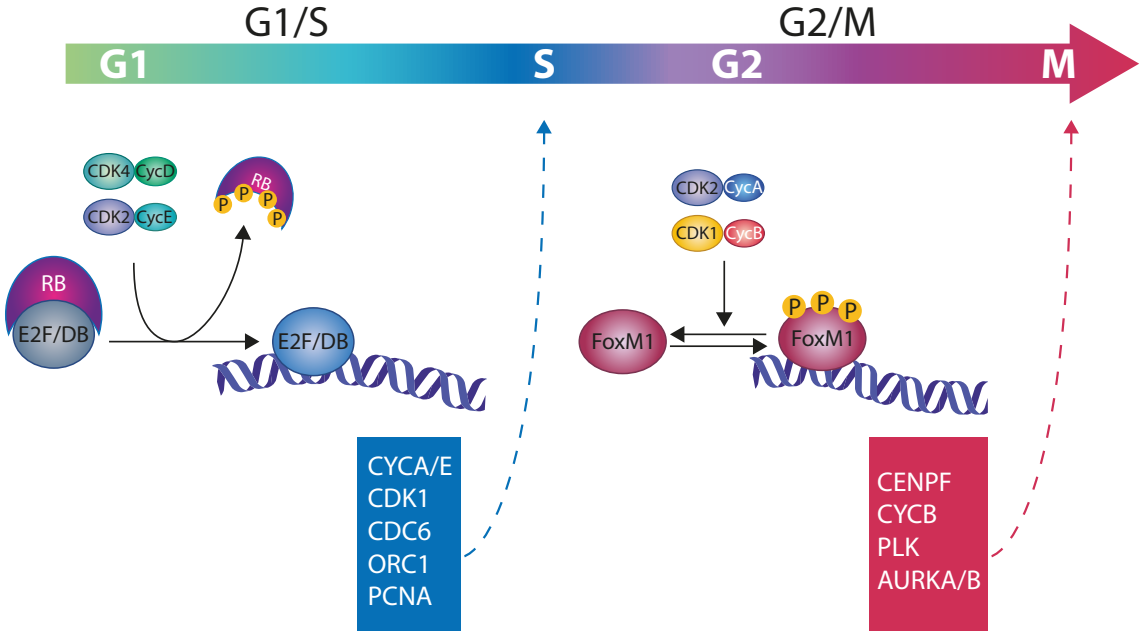


Figure 4: The transcriptional waves of the cell cycle. In the transition to G1/S, cyclins and CDKs phosphorylate Retinoblastoma (Rb) and prevent its inhibition of E2F transcription factors. In turn, this leads to an increased gene expression of S cyclins, proteins of DNA replication, and other genes with a role in the following steps (in turquoise). In G2/M, cyclins and CDKs activate the FoxM1 transcription factor, resulting in the upregulation of M-phase genes (in red).

d) A multi-layered, regulatory interplay: Checkpoints

A robust control of the cell cycle progression is mainly achieved by the previously discussed layers of regulation. In turn, these layers of regulation are further controlled by additional regulatory proteins outside the cyclin-CDK control system, by the effector proteins themselves via feedback loops, by integrating third-party signaling such as controlling cell size or mitogenic extracellular factors, and by cross-talking between each layer of regulation. Moreover, all of these regulatory mechanisms not only control the events in their own cell cycle phase, but also those of other phases.

Further modulation of CDK-cyclin complex activity occurs by full activation or inhibition. Full activity is achieved by CDK-Activating kinases (CAKs), as cyclin binding alone cannot fully activate CDKs. These enzymes phosphorylate a residue adjacent to the kinase active site that further enhances catalytic activity of CDKs, a modification that is not reversible. Kinases in charge of activating CDKs not necessarily bear homology across species, with animals using another CDK-cyclin in a larger complex and yeasts a monomeric protein. To prevent overlap between different complexes, CDKs and cyclins of a certain phase are inactive in the rest. Because dephosphorylation by CAKs is not reversible, CDK activity is downregulated by CDK-Inhibitor Kinases (CIKs), and CDK-inhibitor proteins (CKIs). Inhibitor kinases can phosphorylate kinases in different residues than CAKs to drive them in an inactive state (Morgan, 2007). Inhibitor proteins bind and inactivate cyclin-CDK complexes: one important example are the proteins p21 and p16, from the INK4 family stabilising G1 phase and preventing re-entry into cell cycle (Beach et al., 1993; Polyak et al., 1994; Wade Harper et al., 1993; Xiong et al., 1992). Another example is the protein Wee1, a kinase that prevents entry into mitosis by inhibiting mitotic CDKs (Den Haese et al., 1995) (Figure 5). Cyclin levels are in turn lowered by protein degradation. A ubiquitin-ligase macromolecular complex named the Anaphase-promoting complex (APC) degrades different targets of the cell cycle depending on different regulatory subunits. S and M cyclins and securin, among other proteins, are degraded in the presence of Cdc20 subunit at the end of M phase. Securin is the first step in the cascade to degrade cohesins and promote sister chromatid separation (Cohen-Fix et al., 1996; Shirayama et al., 1999). M CDKs are degraded when APC binds Cdh1 during G1 until before entering to a new cycle (Visintin et al., 1997).

Some additional mechanisms of regulation are only now being investigated, such as tridimensional chromatin architecture and epigenetic regulation across different cell cycle

phases. Concomitant to the unfolding of chromatin following DNA replication, early-S phase chromatin shows a lower degree of insulation in topologically-associated domains compared to later-S, mitosis or G1, where the DNA is tightly packed and later expanded (Nagano et al., 2013). These changes in chromatin are accompanied by active and inactive epigenetic marks, respectively (Pope et al., 2014). In this deterministic scenario of cell cycle-regulated chromosomal dynamics, the cell still manages to integrate regulatory programs -like responses to stimuli or cell fate specification- that often imply genome reorganization and epigenetic regulation, but how this is done is still beyond full comprehension.

Cell cycle regulation is more than independent layers of regulations: First, these different mechanisms interact with each other. The periodic transcriptional program is regulated by the cyclin-cdk control and vice versa: upon entry into S phase, G1 cyclin-CDK complexes phosphorylate the repressor protein Retinoblastoma (Rb) to release the transcription factor E2F and trigger the upregulation of genes in S phase (Fig. 4). In turn, expression level of mitotic cyclins is usually regulated by the transcription factors controlling the mitotic transcriptional wave (Costa, 2005; Pramila et al., 2006) (Fig. 4). Secondly, cell cycle regulation occurs across time between phase-specific regulators: some can inhibit those of the previous phase and upregulate those of the next phase. For example, cyclin-CDK complexes in G1 increase the expression of S cyclins and later M cyclins (Morgan, 2007). Other examples are CKIs Kip/Cip, which repress G1 cyclin/CDKs but are required to form active G1/S cyclin/CDK complexes (Polyak et al., 1994). The APC keeps active during G1 by binding the Cdh1 subunit, preventing early activation of mitotic cyclins and thus promoting unidirectional progression through the cell cycle (Fang et al., 1998; Sudo et al., 2001; Visintin et al., 1997).

Lastly, the cell cycle is not an isolated process occurring indifferent to the surrounding context. Different ways to integrate input from other processes in the regulation of the cell cycle include cell size and nutrient monitoring systems, such as Cln3 or PFK2 in yeasts (Newcomb et al., 2003), or the Cyclin D-CDK4,6- mediated response to mitogens in animal cells.

The resulting regulatory interplay of cyclin-CDK control, transcriptional regulation, temporal regulatory feedback, and third-party modulation, allows the cell to monitor both internal and external cues over time in the so-called checkpoints. Only when all the requirements of a certain checkpoint are satisfied will the cell decide to progress through:

before committing into DNA replication and mitosis, it is required to ensure the cell is large enough to give rise to two newborn cells, that the DNA is not damaged before DNA replication, and that the environment is suitable. The regulatory elements controlling these aspects constitute the G1/S checkpoint. The cell will not proceed into mitosis until all chromosomes are attached to a spindle tip, and until the two centrosomes are located in the plane of division, constituting the spindle checkpoint. Sister chromatids will not separate until all the chromosomes have been replicated and they are positioned in the metaphase plate attached to the mitotic spindle, constituting the mitotic checkpoint. Overall, checkpoints reflect a reinforced selective pressure for cells to follow a sequential, logical order of events strictly in response to very specific conditions.

1.4 Adaptations of the cell cycle in different life situations

Cell reproduction is so pivotal and deeply rooted in the life cycles of all species that it is almost logical to think it is different for each organism. However, a broad glance across different eukaryotic lineages reveals as many differences as similarities, the latter sometimes convergent. It is neither easy nor practical to describe the cell cycle regulation of different organisms without a comparative approach. For this reason, the writing of cell cycle features below is inevitably biased towards the study of classical model systems; in particular, and pertinent to our topic of study in this thesis, a focus is done into comparisons from the perspective of the cell cycle in cells from multicellular systems, namely animals.

a) The prokaryotic cell cycle

Although the work in this thesis focuses on cell cycle regulation in eukaryotes, prokaryotes also follow progression through a cell cycle. There are two main groups of prokaryotes: Bacteria and Archaea (Madigan et al., 2008). They share some similarities due to their prokaryotic organization, such as the genome encoded in one single circular molecule and the presence of a cell wall, but differences in multiple aspects place Archaea and Bacteria as very distantly related lineages. Cell cycle in prokaryotes is one of those different aspects.

The most studied bacterial cell cycle is that of *Escherichia coli* (Dewachter et al., 2018; Donachie, 1993), where a cell size control mechanism triggers the initiation of DNA

replication upon growth to a critical mass. A protein named DnaA binds to the only origin of replication in the chromosome, oriC, and opens the different strands. Unlike the linear chromosomes in eukaryotes, bacterial chromosomes are circular and end up physically interlinked after DNA replication, their decatenation mediated by topoisomerases. Simultaneously to DNA replication, chromosomes start segregating by yet unclear mechanisms, with both passive (entropy) and active (protein activity) forces putatively involved (Graumann, 2014). Septation occurs at mid cell and is coordinated with chromosome segregation (Donczew et al., 2016). The Z ring, a dynamic ring of tubulin-homologue polymers named FtsZ, acts as a scaffold to recruit numerous proteins involved in the generation of constriction forces that ultimately divide the cell in two.

The cell cycle in Archaea, however, is more similar to that of eukaryotes. Archaeal cell cycle phases are similar to the described above and spend most of their time in their equivalent of G2 (Lindås and Bernander, 2013). Many mechanisms and effector proteins are conserved between archaea and eukaryotes, including the cytoskeletal and replication machinery (Orc Cdc6, Mcm); Archaea can have multiple replication origins, in contrast to Bacteria (Wu et al., 2014). These similarities reflect the evolutionary origin of eukaryotes within the Archaea division (Eme et al., 2017). Interestingly, serine-threonine kinases have been found to highly oscillate during the archaeal cell cycle, hinting to the fact that maybe regulators similar to CDKs may also be present in Archaea (Lundgren and Bernander, 2007).

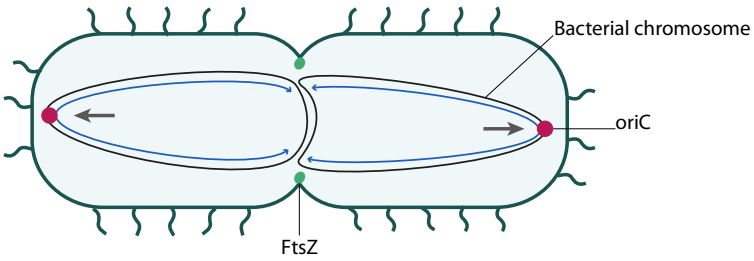


Figure 5: prokaryotes have a different cell cycle. A simplified model of cell division in *Escherichia coli* by binary fission. Newly synthetized DNA is shown in blue. Chromosomes (black) are pulled apart from the origin of replication (oriC) as DNA replication concludes, and a cell wall constriction takes place at midcell in the Z ring (FtsZ).

b) The cell cycle in unicellular eukaryotes: Yeasts as a model system

Yeasts have historically been a powerful model system to understand cell cycle regulation in eukaryotes with all the advantages of a unicellular organism, such as a rapid generation time or the ability to do quick, large scale mutant screenings. Specifically, two species have been largely characterized at the cell cycle level: the budding yeast *Saccharomyces cerevisiae*, and the fission yeast *Schizosaccharomyces pombe*.

Yeast cell cycle structure slightly differs from the arrangement described in previous sections. While DNA replication and mitosis still occur in separate phases, there is no G2 in *S. cerevisiae* phase because budding begins when the cell starts replicating the DNA and chromosomes can be segregated immediately after DNA replication. Recent findings hint that these processes are so overlapped that DNA replication even finishes in anaphase (Ivanova et al., 2018). Moreover, full cell separation occurs already into the next cycle (Morgan, 2007). In *S. pombe*, the cell has already passed the cell size threshold to enter a new cell cycle before completing former cytokinesis; thus, DNA replication starts as the cell septates (Hagan et al., 2016; Mitchison and Creanor, 1971) (Fig. 6). In addition, yeast mitosis occurs without the disassembly of the nuclear membrane, and the machinery used to separate chromosomes, named the spindle pole body, is different from the centrosome (Kilmartin, 2014). Entry into the cell cycle is controlled by external cues, such as the presence/abscence of mating pheromones which can trigger arrest or progression through the cycle (Strickfaden et al., 2007), or intracellular cues such as cell size regulated by dillution of the repressor Whi5 (Schmoller et al., 2015).

Albeit present in all eukaryotes, the cyclin and CDK gene families have undergone independent expansions and subfunctionalization in every major eukaryotic lineage, including opisthokonts --the branch where animals and fungi are found, among others (Figure 1A) (Cao et al., 2014; Medina et al., 2016). For example, in *S. cerevisiae*, one single CDK sequentially binds to nine cyclins in three temporal waves (Bloom and Cross, 2007; Mendenhall and Hodge, 1998): Cln1-2 are expressed in G1 and mark the commitment to a new cycle (Cross and Tinkelenberg, 1991; Nasmyth and Dirick, 1991; Richardson et al., 1989; Skotheim et al., 2008), Clb5,6 promote DNA replication at S phase (Epstein and Cross, 1992; Schwob and Nasmyth, 1993), and Clb 1-4 drive progression through mitosis. The fission yeast *S. pombe* has a single CDK that also binds different cyclins at G1,S, and M: Cig1,2 drive progression through G1 and S phase

(Martín-Castellanos et al., 2000; Mondesert et al., 2015), and Cdc13 drives progression through mitosis (Booher and Beach, 1988; Booher et al., 1989). However, a single CDK-Cyclin complex can drive progression through the entire cell cycle in this species (Coudreuse and Nurse, 2010). CAKs, the kinases in charge of fully activating CDKs, are partially conserved with their animals counterparts (Morgan, 2007).

Transcriptional regulation of the G1/S transition in yeasts is driven by SBF and MBF, two transcription factor complexes that are exclusive to Fungi and bear no homology to the paneukaryotic transcription factor E2F (Iyer et al., 2001). Recent findings show that these transcription factors were acquired through lateral gene transfer from a virus during fungal evolution (Medina et al., 2016). In addition, other transcription factors in *S. cerevisiae* are the mitotic Fkh2 (Gefeng et al., 2000; Reynolds et al., 2003) or Hcm, a Forkhead-class transcription factor controlling progression through G2 and mitosis (Pramila et al., 2006). Some of these transcription factors are coordinated within the same network in budding yeast, whereas in fission yeast they operate in parallel (Rustici et al., 2004). The cell cycle transcriptional program has also been investigated in the fungal pathogens *Cryptococcus neoformans* and *Candida albicans*. Interestingly, only a relatively small set of common genes is cell cycle-regulated in the three species, with many shifting their time of peak expression in the cell cycle (Côte et al., 2009; Kelliher et al., 2016). Thus, periodic transcriptional programs across yeasts are similar only at the surface level, as many of their periodic genes obey to the production of secondary metabolites or the two entirely different mechanisms of cell division that are, in the case of *S. cerevisiae* and *S. pombe*, bud formation and cell fission (Pramila et al., 2006; Rustici et al., 2004).

Overall, it has become increasingly clear that many otherwise conserved cell cycle regulators have been lost and independently evolved in the fungal lineage compared to the ancestral opisthokont (Cooper, 2006; Cross et al., 2011; Kearsey and Cotterill, 2003; Medina et al., 2016; Rhind and Russell, 2000). This is most dramatic in yeasts, known to have lost and reinvented a substantial number of other cellular components (Gallone et al., 2016; Nguyen et al., 2017; Rhind et al., 2011; Steenwyk et al., 2019).

c) Other unicellular eukaryotes

Social amoebae

Amoebozoa is a group of unicellular amoeboid eukaryotes that, together with opisthokonts and other lineages, form the clade Amorphea (Figure 1). Within Amoebozoa, cell cycle regulation has also been investigated in the so-called social amoebae: *Dictyostelium discoideum* and *Physarum polycephalum*.

D. discoideum is an amoeba that grows in unicellular populations until starvation and other signals induce the aggregation of cells into a “slug”, a superorganism with the ability to migrate (Robertson and Grutsch, 1981). After settling, the slug transforms into a sessile structure with a stalk and a fruiting body, from which spores of single cells are released. The fate of a cell within the whole entity (slug head, tail, stalk, fruiting body…) is determined by their cell cycle phase when aggregation occurs (Gomer and Firtel, 1987). In cell cultures of *Dictyostelium*, cells spend a long time in G2 phase; S phase and M are short in proportion, and there is no detectable G1 (Weijer et al., 1984). Other than that, it has a conventional cell cycle toolkit with one CDK, different cyclins of B/C/H/L/T/Y types (Cao et al., 2014), a recognizable APC, aurora and polo kinases, and a retinoblastoma ortholog involved in the regulation of typical S and M phase genes (Strasser et al., 2012). There is little known about temporal waves of gene expression, or how it regulates entry into the cell cycle.

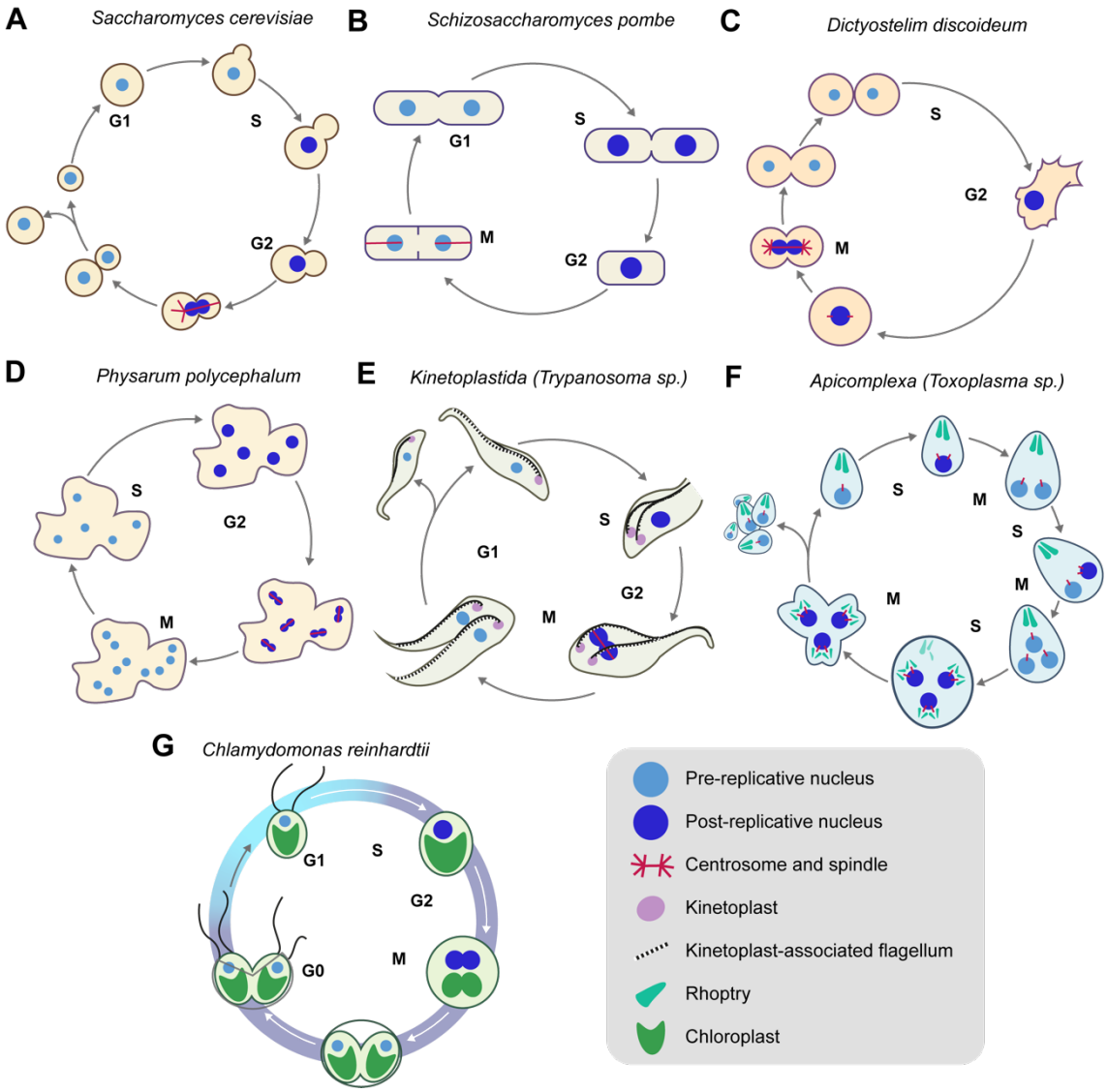


Figure 6: The diversity of cell cycles across unicellular eukaryotes.

In *P. polycephalum*, the vegetative phase of this species is a plasmodium, a syncytium where all the nuclei are surrounded by a common plasma membrane forming a giant, multinucleated cell (Guttes and Guttes, 1964). Nuclei within a plasmodium divide synchronously by accumulation of a cytosolic spreading factor that triggers mitosis (Guttes et al., 1961; Rusch et al., 1966; Sachsenmaier et al., 1972)). There are documented evidences accounting for a mass/size controller acting in G2 (Dove et al., 1986); thus, the checkpoint triggers nuclear division followed by immediate DNA synthesis, leaving no G1 phase such as in *D. discoideum*. Concomitantly, cyclin B1 accumulates during G2 phase and likely mediates progression through mitosis (Rajan et al., 2014). There are waves of expression of important genes for DNA replication such as histones (Dove et al., 1986). DNA polymerase peaks at G2 matching the readiness

of the replication machinery upon chromosome segregation, and tubulin peaks at mitosis, likely because they lack a inactive tubulin stock as typical of mammalian cells (Dove et al., 1986).

Kinetoplastids and apicomplexans: parasitic and pathogenic protists

Because of their impact in human and animal health, two eukaryotic groups where the cell cycle has been heavily studied are kinetoplastids and apicomplexans.

Kinetoplastids derive from a group of photosynthetic organisms within Excavata named euglenids and have a kinetoplast, a unique organelle harboring a large amount of DNA in association to the basal body of the flagellum. Some of them are free living, but in the group *Trypanosoma*tids we find the pathogens *Trypanosoma* and *Leishmania*. Kinetoplastids have their own set of ten independently expanded cyclins, with three of them resembling B-type cyclins of animals and other eukaryotes. Similarly, CDKs have evolved independently into CRKs. It is likely that CRK activity is regulated by divergent activators and inhibitors, as many CDK residues key to canonical CAK/CKI regulation are different in CRKs; in parallel to this, only a limited number of these activators and inhibitors are conserved with other eukaryotes (Li, 2012). Cell cycle progression is mediated by CRKs and cyclins: CRK1,2 and CYC2,4,5,7 are required for progression through G1 and G1/S (Gourguechon and Wang, 2009; Hammarton et al., 2004; Li and Wang, 2003) and CRK3 (homolog of animal mitotic CDK), CRK9, and CYC2,6 are known to control progression through mitosis (Gourguechon and Wang, 2009; Hammarton et al., 2004, 2003). In S phase, DNA replication origins are recognized by a highly derived set of ORC proteins, and Kinetoplast DNA is replicated by an independent machinery from the nuclear one (De Melo Godoy et al., 2009; Tiengwe et al., 2012). How chromosomes segregate is still not very clear, since less than the required amount of kinetochores per chromosome are formed (Tan et al., 2005). Interestingly, cell cycle control is deployed differently between *Trypanosoma* life stages, with bloodstream cells lacking a cytokinesis checkpoint, contrary to procyclic cells (Wheeler et al., 2013).

Apicomplexa encompasses a group of single-celled, obligate intracellular parasites distantly related to a group of free living unicellular algae within alveolates (Woo et al., 2015). Apicomplexans differ in host and life cycles but share common adaptations to parasitic lifestyle such as the apical complex, an organelle made of cytoskeletal and secretory (White and Suvorova, 2018). Among these adaptations, they can decide to undergo different types of cell cycles, concomitant to the parasitic lifestyle: from budding

to fission and from one replication to multiple DNA replications (Ferguson et al., 2008). The molecular basis of the cell cycle is formed by a mixture of conserved and newly evolved proteins. As *Trypanosoma*tids, Apicomplexa have a derived set CDKs also named CRKs. G1/S transition is regulated by the ApiAP2 family of transcription factors, instead of E2F/Rb. This family was acquired through lateral gene transfer from an algal endosymbiont in the photosynthetic ancestor (White and Suvorova, 2018). DNA replication machinery is conventional, but regulated by a novel kinase (Matthews et al., 2018). No G2/M checkpoint exists as no G2 phase is detected in apicomplexans. In mitosis, chromosome segregation and the spindle assembly checkpoint are regulated by aurora kinases and the APC, but polo-like kinases have been lost and another CRK controls exit from mitosis (Ganter et al., 2017; Suvorova et al., 2013). Interestingly, apicomplexans can decouple karyokinesis from cytokinesis during endopolygeny by using two independent microtubule organizing centers, in order to produce large offsprings in their infection cycles (Courjol and Gissot, 2018). This is controlled by a newly evolved CRK that is able to trigger chromosome segregation only in nuclei in a post-replicative state; the mechanisms by which these are distinguished from those in S phase are still unknown although local regulation has been suggested (Dorin-Semblat et al., 2013; Francia and Striepen, 2014; Naumov et al., 2017).

Overall, because of their condition as unicellular parasites, the cell cycle control in *Trypanosoma*tids and apicomplexans is tightly related to developmental control, and they have evolved different ways to regulate cell cycle control in compensation for the gene losses concomitant to their parasitic lifestyles (Li, 2012; White and Suvorova, 2018).

Photosynthetic unicellular eukaryotes

In *Chlamydomonas reinhardtii*, a unicellular green alga closely related to multicellular volvocales algae (Figure 1A,E) (Herron, 2016), progression through the cell cycle is coupled to the circadian clock (Zones et al., 2015). Cells proceed synchronously through G1 during the day and undergo multiple rounds of DNA synthesis and mitosis during the first night hours. Newborn cells remain in G0 until dawn, where they start growing in mass again. Due to this deployment of the cell cycle, most of Chlamydomonas genes are periodically expressed throughout day/night shifts, including those encoding photosystems and biosynthesis of pigments proteins ((Zones et al., 2015). In this program, most cell cycle-related genes peak before (DNA replication machinery) or after

(mitosis) dusk, including many conserved cell cycle regulators such as the APC, E2F1, MAT, CKS1, cyclins, and CDKs. The latter two have also undergone independent expansions in Archaeplastida (the group comprising green algae and plants) (Gutierrez, 2009).

Dinoflagellates, another group of photosynthetic eukaryotes with remarkable impacts in marine ecology (Hackett et al., 2004; Wang, 2008), follow a similar pattern. They complete a typically-phased cell cycle in a 24 hour period, with G1 during the day, S phase in the first hours of night, and G2/M during the rest of the night (Bhaud et al., 2000; Wang et al., 2013); these cell cycles can be affected by perturbations in the environment such as light and nutrient (Olson and Chisholm, 1986; Taroncher-Oldenburg et al., 1999). Phylogenetic analyses have shown that dinoflagellates have cyclins and CDKs from the same families as animal S and M cyclins and CDKs (Morse et al., 2016). Among these, homologs of conventional G1, G1/S, and M cyclins have been documented to peak at their respective phases (Barbier et al., 2003; Shi et al., 2017; Zhuang et al., 2013). Although protein levels of histones and proliferation markers during dinoflagellate cell cycle match with cell cycle dynamics previously described in other species (Brunelle and Van Dolah, 2011; Wang et al., 2013), little is known about global cell cycle gene expression in these organisms.

From the study of all these unicellular eukaryotes, we have learnt that cell cycle in unicellular species is tightly coupled to other fundamental processes of their lifestyles, such as complex development life cycles, circadian rhythms, mating, or metabolism. Secondly, we already observe large differences in cell cycle deployment in a not-so-large number of examples that only represents a small percentage of the currently estimated eukaryotic diversity (Del Campo et al., 2014). While cell cycle regulation has been studied in a selected number of eukaryotes for different reasons (like being pathogens, because they are model systems, or because they have an ecological impact), their differences in cell cycle regulatory programs point to a larger unknown diversity of cell cycle regulation awaiting to be unraveled.

d) Multicellular organisms: animals and plants

Animals and plants belong to opisthokonts and archaeplastids, respectively, two lineages of eukaryotes that are very distantly related. Despite this, they share many similarities due to their embryogenic multicellular lifestyles. In contrast to unicellular organisms, animal and plant cells grow and proliferate under certain conditions according to the needs of the whole collective, such as regeneration or embryonic development (Figure 1D). This occurs by integrating various signals from the multicellular environment in which they are found: in both lineages, G1 cyclins control entry into cell cycle in response to extracellular mitogenic and hormonal signals ((Matsushime et al., 1994, 1991); reviewed by (Sherr, 1995) and (Klein and Assoian, 2008); (Gutierrez, 2009)). Animals and plants are formed by many different cell types, and cell cycle regulation varies across them, often integrating with the ontogenetic plan of the cell type.

Many cell cycle genes are conserved in animals and plants, such as the E2F-Rb network, PCNA proteins, Aurora kinases, the APC, or the cyclin and CDK subfamilies. Compared to other model systems such as yeasts, and concomitant to historic events of genome expansions and duplications in animals and plants (Paps and Holland, 2018; Rejlová et al., 2019; Richter et al., 2018), some of these gene families of cell cycle regulators have undergone extensive diversification and specification, acquiring additional layers of regulation to time their activity (Menges et al., 2003; Rustici et al., 2004; Whitfield et al., 2002).

In animals, progression through the cell cycle is regulated by an elaborate system of cyclin-CDK interactions where specific partners activate downstream effectors at different stages of the cell cycle. While many metazoan groups have expanded cell cycle cyclins and CDKs differently, members of cyclin A, B, D, E, and CDK 1, 2, 4 and 6 subfamilies have been found in representative species of all metazoans (Cao et al., 2014; Gunbin et al., 2011; Liu and Kipreos, 2000). Both cyclins and CDKs are transcriptionally and post-translationally regulated, showing specific timings of expression, subcellular localizations, and affinities for substrates (Morgan et al., 2007; review specific cyclin b and stuff). Cyclin D-CDK4,6 complexes control entry into the cell cycle in response to mitogenic factors ((Matsushime et al., 1994, 1991); reviewed by (Sherr, 1995) and (Klein and Assoian, 2008)). The G1/S transition is driven by the Cyclin E/CDK2 complex, and progression through S phase is controlled by the Cyclin A/CDK2 complex (Cook et al., 2002; Coverley et al., 2002). Lastly, cyclin B/CDK1 drive

completion of mitosis (Pomerening et al., 2005, 2003). Because some of these interactions are shared (although not promiscuous) and occasionally differ across cell types, cyclin and CDKs are thought to retain the ability to bind to other members aside from their canonical partners. This results in multiple mechanisms of redundancy, useful in the case of an abnormal scenario where one of the components fails. For example, animal cell cycle can still work in the absence of a number of G1 and S CDK partners, suggesting only CDK1 is essential to drive progression through the cell cycle (Santamaría et al., 2007).

The CDK repertoire found in plants comprises six subfamilies, named CDK A-F. Two of these, CDKA and CDKB, bind to G1/S and M cyclins and drive cell cycle progression, with CDKA (related to human CDK1 (Medina et al., 2016)) binding to all cyclins much like in yeast (Gutierrez, 2009). Most of C to F-type CDKs are involved in the cell cycle one way or another, some acting as CDK activators (CAK) (Gutierrez, 2009). Plant cyclins are representative members of A and B subfamilies, also present in animals and other eukaryotes (Cao et al., 2014; Medina et al., 2016). Plants “D-type” cyclins, which are ortholog to the Cyclin G family in other eukaryotes (Medina et al., 2016), interact with RB to promote entry into cell cycle, and B-type cyclins regulate progression through mitosis and are degraded at M phase by the APC.

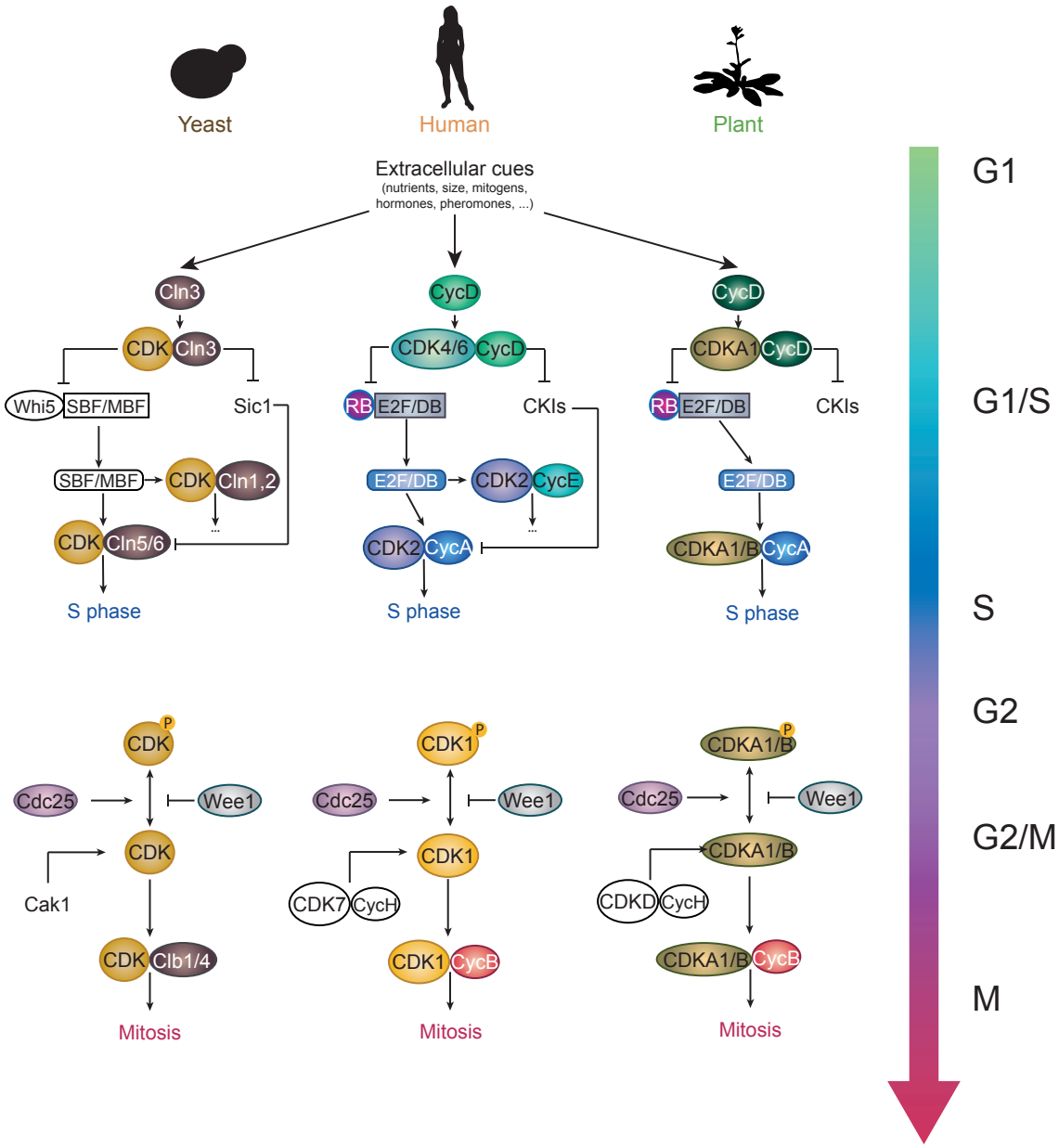


Figure 7: The main cell cycle checkpoints and regulators across species. Generally, cyclin-CDK complexes indirectly upregulate transcription factors by inhibition of repressors. These transcription factors may, in turn, promote the expression of cyclin-CDKs from later stages. In parallel, cyclins and CDK also regulate (and are regulated by) different activators and inhibitors, such as cdc25 or wee1, which ensure proper transition through cell cycle stages. Notice the different degree of conservation in the regulatory networks between species: CDKs and transcription factors of the same color are phylogenetically related. Cyclins of the same color hue play a part in the cell cycle phase according to the color on the right axis.

Gene expression during the cell cycle has been characterized mainly in human and mouse cells (Bar-Joseph et al., 2008; Cho et al., 2001; Grant et al., 2013; Ishida et al., 2001; Whitfield et al., 2002), and in cell suspension cultures of *Arabidopsis thaliana* and *Nicotiana benthamiana* (Breyne et al., 2002; Menges et al., 2003; Menges and Murray, 2002). Similar to other organisms, there are two main temporal waves of gene expression, one at the G1/S transition and another at G2 and M. The G1/S wave is controlled in both lineages by the same E2F/Rb network in both animals and plants, which is essentially conserved (Gutierrez, 2009). E2Fs comprise a gene family of both transcriptional factors which activate when binding to other proteins from the DP family (Helin et al., 1993). Animals E2F1,2,3 (most notably E2F1) and plants E2Fa,b,c are known to activate gene expression (as in other eukaryotes described above), and E2F4,5,6,7 are transcriptional repressors (Litovchick et al., 2007). E2F1 regulates the expression of NFY (Caretti et al., 2003; Zhu et al., 2004), a multimeric transcription factor that can act as an activator or repressor depending on its cofactors and thus binds to the promoters of cell cycle genes from different temporal waves. E2F1 also controls b-Myb, a transcription factor that activates expression of genes at G2/M (Zhu et al., 2004). Another transcription factor of G2/M genes is FoxM1 (Costa, 2005; Laoukili et al., 2005), from the same family as yeast Hcm1 (Pramila et al., 2006). b-Myb and FoxM1 are recruited sequentially by the MuvB complex, a major regulator of G2/M gene expression in mitosis (Sadasivam et al., 2012). Overall, the network of transcription follows a sequential interaction of E2F, NFY, Myb, MuvB and FoxM1. Orthologs of these proteins have been found in many eukaryote species, including plants: *Arabidopsis thaliana* uses different Muv-B complexes to regulate G2/M progression (Fischer and DeCaprio, 2015; Kobayashi et al., 2015a), and transcription factors of the Myb family also controls expression of G2/M genes in *Nicotiana tabacum* (Araki et al., 2004; Ito et al., 2001).

Several of these studies identify periodic genes related to programmed cell death, differentiation, and development, showing the specification of the cell cycle towards multicellular programs (Bar-Joseph et al., 2008; Cho et al., 2001; Dominguez et al., 2016; Grant et al., 2013; Gutierrez, 2009; Iyer et al., 1999; Menges et al., 2003; Whitfield et al., 2002). Animal and plant periodic transcriptional programs are nonetheless very divergent; one example is cell plate-mediated cell wall deposition in plants to septate daughter cells, a mechanism of cell division lacking homology to the actomyosin ring in animal cytokinesis (Guertin et al., 2002; Gutierrez, 2009). Differences occur also within the same species: of hundreds of periodically expressed genes detected in different human cell types, only a fraction is conserved across them (Giotti et al., 2017; Grant et

al., 2013), the others related to the cell type physiology or byproducts of cancerous deregulation in the immortalized cell lines.

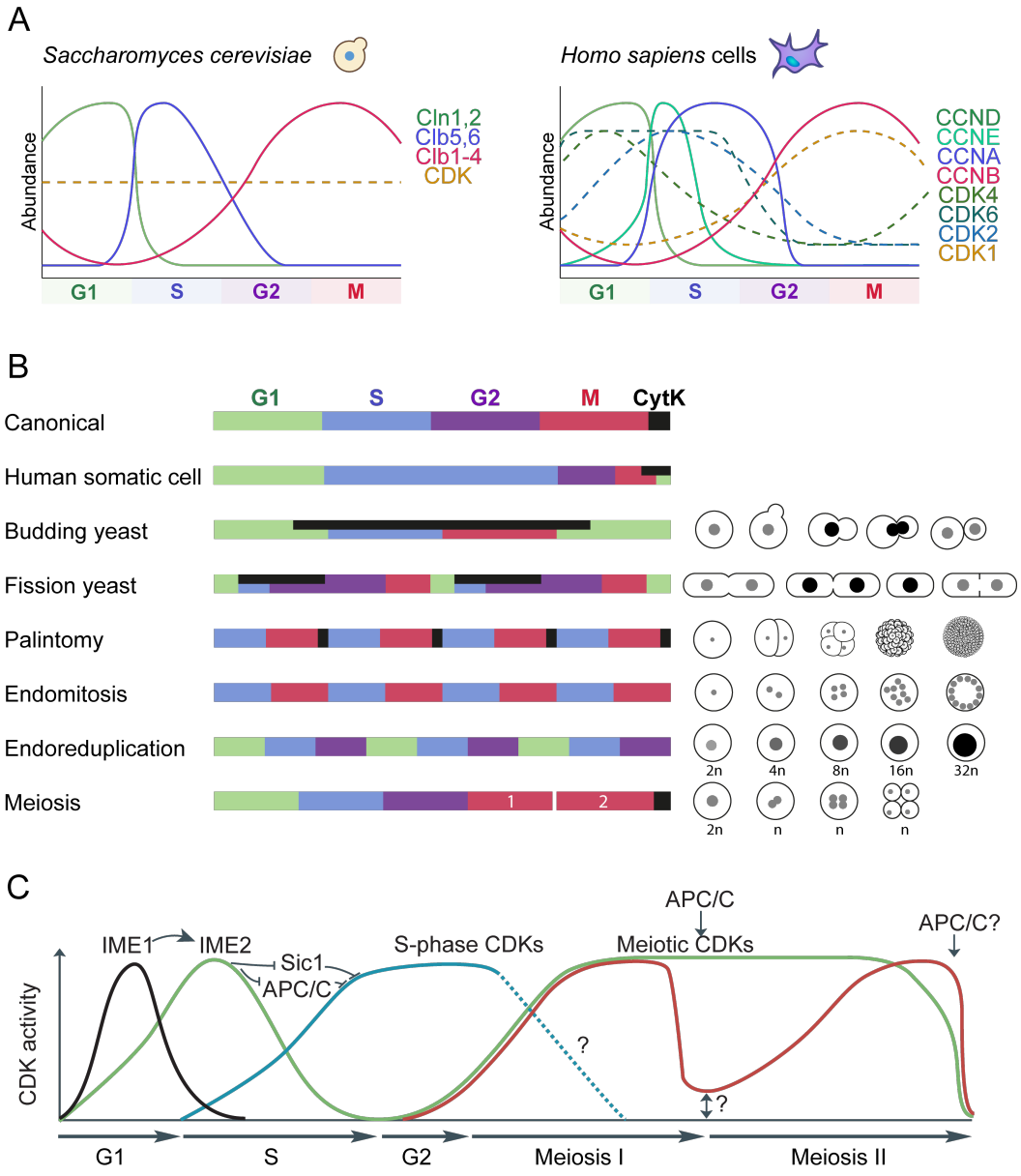


Figure 8: structure and regulation of the cell cycle in unicellular and multicellular organisms. A: temporal dynamics of the core cell cycle system (cyclins and CDKs) in two model systems: the budding yeast and somatic human cells. Continuous lines represent cyclins, and dashed lines represent CDKs. Note that yeast CDK remains transcriptionally stable during the cell cycle, while multiple CDKs oscillate in animal cells. Color code represents different stages of the cell cycle. B: different structures of the cell cycle between different organisms and cell types. Phases are color-coded as indicated for the canonical structure. On the right, visual representation of the resulting phenotypes of each cycle. Gray nuclei indicate prior to DNA replication, black nuclei indicate post-DNA replication. C: the meiotic cell cycle. Adapted from (Marston and Amon, 2004).

Cell cycle regulation is linked to embryogenesis

Animals and plants show embryogenic multicellularity: cells arrange following a developmental program that culminates in many different cell types coexisting and filling diverse roles in a cohesive body plan. These developmental programs are different both between and within animals and plants. During animal development and organogenesis, many cell types specialize and never enter into the cell cycle again, while the ability to proliferate is retained in a very small number of cells committed to give rise to specific lineages of cell types. Conversely, speciallized cell-types in plants can often revert to a pluripotential stage and become proliferative, allowing plant organs to keep developing after embryogenesis. Because of these differences in developmental plans and commitment to proliferate, cell cycle regulation often deviates from the canonical stage structure during development and between cell types within a multicellular organism.

For example, in early stages of animal development, the fertilized egg rapidly divides by palintomic division into thousands of small daughter cells without an increase in cell size, in order to generate enough cells for the newborn organism to fend (Morgan, 2007). In the vertebrate *Xenopus laevis*, this occurs by cells alternating between S and M while skipping G1 and G2 phases; in other systems such as insect embryos, nuclei multiply without cell division by skipping cytokinesis, forming a coenocyte, a giant multinucleated cell, after which cellularization takes place (Farrell and O'Farrell, 2014) (Fig. 8C). This rapid consecution of events is highly coordinated by molecular clocks that run autonomously and independent from checkpoints. At this point, timing of the cell cycle relies in the so-called maternal factors, molecules of messenger RNA and proteins inherited by the maternal egg cell, with little to none transcriptional regulation during these stages. Such circumstances lead to a number of differences from the canonical cell cycle regulation, such as the aforementioned high levels of G1/S cyclin E throughout the cell cycle and S phase cyclin A acting as cyclin B (Morgan, 2007). Later in the ontogeny, the cyclin/CDK system and the periodic transcriptional program take control of the cell cycle in most somatic cells. Regarding different cell types and tissues, some specific cells become polyploid by replicating their DNA without chromosome segregation, such as cells in salivary glands of the fruit fly undergoing endoreduplication, or polyploid hepatocytes and megakaryocytes undergoing endomitosis (Ullah et al., 2009). In plants, many different, large cell types become polyploid by endoreduplication, such as ovular cells, root hair cells, or leaf epidermal cells (Joubes and Chevalier, 2000).

Cell cycle regulation is intertwined with wound healing and regeneration. Many animals are capable of regenerate structures (from tissues to whole organs, limbs or body parts) upon damage or amputation. Generally, cells from the surrounding damaged tissues dedifferentiate or re-program to form a blastema, a structure in which cells proliferate and differentiate following a developmental pattern. In mouse ear regeneration, G2-arrested fibroblasts form the blastema and undergo cell division (Heber-Katz et al., 2012); in axolotl (salamander) limb regeneration, multiple cell types contribute to form the blastema by exiting G0 and re-entering into G1 (Gerber et al., 2018). In both cases, the blastema is formed and proliferates in response to axonal and neural factors, suggesting a key role of the nervous system in regeneration through de-differentiation signaling. Overall, regeneration occurs by a context-dependent regulation of cell proliferation in a specific group of cells able to dedifferentiate and re-enter into a developmental state (Nacu and Tanaka, 2011).

e) Cell cycle and cell identity

The identity of a cell, either defined as a life/metabolic/response stage in unicellulars or a cell type undertaking specific roles in multicellulars, is necessarily linked to the fundamental process that is the cell cycle. Acquiring an identity often involves perception of a stimulus in the surrounding environment and transduction of this signal into a response, often a specific program of gene expression driven by a variety of transcription factors. Perception of these signals can virtually occur at any stage of the cell cycle, which offers an additional dimension of regulation of cell identity and differentiation. Examples are widely found across the tree of life: haploid yeasts can switch their mating type every cell cycle by expression of a specific endonuclease in late G1, enabling sexual reproduction and sporulation in cultures from the same progenitor (Haber, 1998). In *Dictyostelium*, cell fate within the multicellular structure is determined by the cell cycle stage at which starvation is sensed: G1 cells commit to pre-spore cells, but S/G2 cells will become pre-stalk (Gomer and Firtel, 1987). In mammalian embryogenesis, endocrine progenitor cells commit to different fates depending on their exposure to differentiation signals at early or late G1; in mouse, cortical progenitor cells adopt different fates depending on the cell cycle phase when detecting extracellular factors. In mouse bone marrow and *Drosophila* nervous system, S-phase progression is required to activate differentiation programs (Pop et al., 2010; Weigmann and Lehner, 1995).

Although not always necessary, there are numerous examples of cell differentiation and reprogramming linked to cell cycle exit in proliferative conditions. During formation of muscle cells, CDKs repress the activity of myogenic transcription factors in immature cells, maintaining their proliferative capacity until the exit of the cell cycle (Guo and Walsh, 1997; Rao et al., 1994; Skapek et al., 1995). In Zebrafish, cardiac regeneration is linked to polo-like kinase 1, a M phase regulator (Jopling et al., 2010; Poss et al., 2002). Examples go on and on, with documented cases in early branching animals such as Cnidaria (Schmid et al., 1988). The link between cell division and cell fate might come in hand to the extensive chromatin remodeling and epigenetic marking that occurs throughout the cell cycle. Studies in pluripotent stem cells (PSCs) have shown that, overall, gene chromatin is initially marked by ambivalent histone marks in early G1, but remain marked by repression marks (enhanced by cyclin D1 in late G1) throughout the rest of the cycle (Singh et al., 2015). During cell division, the DNA is tightly packed and condensed by numerous chromatin remodelers, a reason for which the majority of DNA-binding proteins are excluded from the chromatin. Some remodelers, such as MLL, are however retained and bookmark promoters, leaving genes primed for activation (Lodhi et al., 2014). Developmental genes may thus express in the new cell cycle if appropriate signaling is also active (Singh et al., 2013). This mechanism, named mitotic bookmarking (Fig. 2A), is essential to establish the pluripotent state and makes mitosis a window of opportunity for changes in cell fate, differentiation, and overall identity, by facilitating hyper-dynamic factor exchange and, to a lesser extent, excluding numerous transcriptional repressors of reprogramming factors (Soufi et al., 2012). Another mechanism to establish different fates occurs by asymmetric cell division, by which daughter cells inherit different factors, modifying their molecular background (Bi and Park, 2012; Knoblich, 2008).

f) Sexual reproduction: a different cell cycle

During sexual reproduction, cells from two different mating types combine their genetic material by fusing together, meaning that they restore the normal amount of genetic material. These cells come from a special cell division called meiosis, which yields newborn cells with half the amount of genetic material as the progenitor cell: if the original cell is diploid, four haploid cells will be yielded. The meiotic cell cycle comprises one round of DNA replication followed by two consecutive rounds of cell division, and shares many similarities with a normal cell division cycle (Fig. 8C). Some of the main regulators are the same between mitotic and meiotic cell cycles: in yeasts, cyclin-CDK complexes

controlling S and G2 phases are the same in mitotic and meiotic cell cycles, but one special transcription factor named IME1 (a CDK-like gene) is in charge of triggering meiotic cell cycle in G1 in response to a low amount of nutrients. IME1 upregulates IME2, a G1 CDK-like gene that controls entry into cell cycle and the later execution of meiotic divisions, together with M cyclin-CDK. This latter complex oscillates twice to trigger the two meiotic divisions. Homologous chromosomes are segregated in meiosis I, and sister chromatids are segregated in meiosis II, for which different mechanisms exist across species to ensure the separation in this order (Marston and Amon, 2004).

Meiosis is related to the formation of gametes, and, by extension, is often found in only a very few number of cells within a multicellular organism. Occasionally, progression through the meiotic cell cycle can be stalled during large amounts of time, such as with female ovulation in mammals (Tripathi et al., 2010). Meiosis is also tightly coupled to processes of cell differentiation: in mammals, egg cell precursors yield oocytes and also other types of cells that assist in a variety of functions during development (Schmerler and Wessel, 2011).

g) Cell cycle genes are involved in the acquisition of multicellularity

Multicellularity requires the evolution of mechanisms to keep the cells together. Among the main candidates to be considered are logically genes related to cell cycle and cell division; for example, cells can be kept together after division by mechanisms involving incomplete separation. Another example would be the evolution of mechanisms to time and coordinate cell division in the context of a whole, or to time the expression of multicellularity genes (such as cell-cell adhesion and signalling) to the cell cycle. Interestingly, complexification of the genetic machinery involved in the cell cycle has been linked to the evolution of multicellularity in several cases throughout eukaryotes, most notably green algae and yeasts.

In green algae, *Chlamydomonas*, *Gonium* and *Volvox* form a clade of unicellular, colonial, and complex multicellular lifestyles, respectively. Cyclin D1 underwent expansion in the lineage from *Chlamydomonas* to *Volvox*, with multiple copies in *Volvox* and different paralogs in *Gonium*. The increase and diversification of D-type cyclins has been suggested to play part in the differentiation of cell types in *Volvox* (Hanschen et al., 2016). D-type cyclins interact with Retinoblastoma, whose predicted interacting site with the E2F transcription factor in *Gonium* and *Volvox* is shorter than in *Chlamydomonas*

and lacks key sites for canonical interaction with Rb. Additionally, *Gonium* Rb induces formation of undifferentiated colonies in *Chlamydomonas* cells by a gain of expression of genes related to cell adhesion in G1. These results suggest Retinoblastoma-E2F were co-opted for multicellularity in the volvocine algae.

Multicellularity can evolve when experimentally selecting for larger individuals. Two examples are culturing in presence of predators (Boraas et al., 1998; Herron et al., 2019) and artificial selection for rapid settling in liquid cultures (Ratcliff et al., 2015). In *Chlorella vulgaris*, clonal colonies emerge under 100 generations in the presence of the predator *Ochromonas vallescia*, in an optimal size to escape predation while exposing all cells to the medium for nutrient uptake (Boraas et al., 1998). *Chlamydomonas* is capable of generating similar colonies in the presence of *Paramecium tetraurelia*, although with more variable sizes and putatively linked to a pre-adaptive ability to form transient multicellular structures in other conditions (Herron et al., 2019). In both cases, multicellularity was inherited, meaning there was an underlying genetic mechanism. When analyzing gene expression of *Chlamydomonas* colonies artificially selected for sedimentation, numerous genes related to cell cycle processes and mitosis showed a different temporal up- and down-regulation compared to unicellular *Chlamydomonas* (Herron et al., 2018). The “snowflake yeast”, so-called by its unique reproductive model, is a strain of *S. cerevisiae* artificially evolved under settling selection that grows clonally while keeping daughter cells attached. Genotyping revealed that disruption of a single transcription factor involved in regulation of cytokinesis and bud septation is enough to generate a multicellular phenotype (Ratcliff et al., 2015).

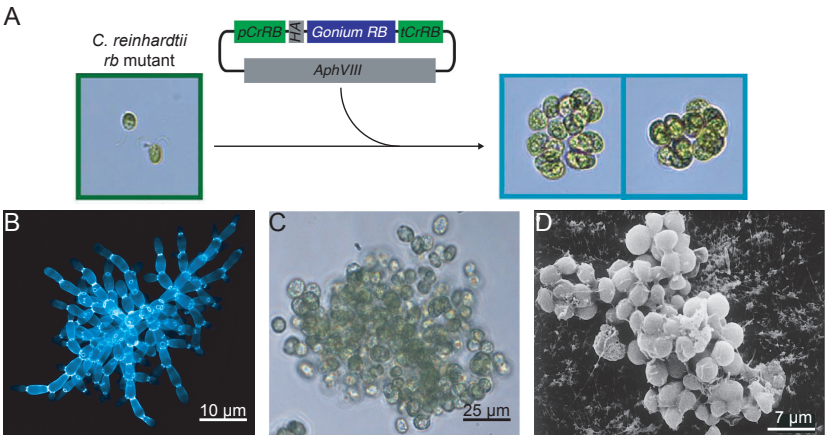


Figure 9: Experimental evolution of multicellularity is linked to cell cycle regulation. A: Introducing a vector carrying the *Gonium pectorale* Retinoblastoma gene in *Chlamydomonas rb* mutants results in the formation of clonal colonies. B: The snowflake yeast, generated by settling selection, is able to grow and reproduce in simple multicellular structures due to mutations in a gene controlling bud fission. C: a strain of *C. reinhardtii* equally generated by settling selection. Peripheral cells seemingly disconnected from the colony are joined by extracellular matrix. D: a colony of *Chlorella vulgaris* evolved by the addition of a phagotroph predator in the culture medium. Adapted from (Boraas et al., 1998; Hanschen et al., 2016; Ratcliff et al., 2015, 2013).

Interestingly, snowflake yeast and other clonal green algae show a developmental model where each propagule, either unicellular or multicellular, forcefully goes through a single-cell genetic bottleneck, thus overcoming the main problem of in-group selection hinted earlier in this chapter (Herron et al., 2018; Ratcliff et al., 2015; Szathmary and Smith, 1995). These cases illustrate how changes in the regulation of cell cycle-related genes can work as the underlying mechanism driving multicellularity.

1.5 Cell cycle evolution: state of the art and paving new avenues

The cell cycle plays an essential part in arguably every aspect of eukaryote life. Cells from all species, regardless of their position in the tree of life, need to attend and solve the same problems related to cell reproduction that are replicating and segregating the DNA. It does not seem strange, then, that the basic components of the same machinery that allows cells to move, change shape, and transport things within them, are the same effectors that mend the problem of segregating chromosomes during cell division, and are thus well conserved. Actin, myosin, tubulin, kinesins and dyneins, among other cytoskeletal components, play important roles in both scenarios and are regulated by the cell cycle machinery. Studies in multiple different eukaryotes have shown that progressing through the cell cycle is essentially governed by the same regulatory core: 1) the cyclin-CDK system, 2) a periodic transcriptional program under the control of mainly E2F transcription factors, and 3) a number of regulators of the former two, including but not limited to kinases and ubiquitin-ligases. Subunits and members of these machineries are widely found across eukaryotes, even in those for which we only have genomic data: such are the cases of the multimeric, intricate anaphase-promoting complex (APC) (Eme et al., 2011), the heavily expanded cyclins and CDKs, the ubiquitous presence of E2F and forkhead transcription factors from plants to animals together with the repressor Rb (de Mendoza and Sebé-Pedrós, 2019; Medina et al., 2016), or diverse proteins such as CAKs, CIKs, Aurora- and polo-like kinases (Archambault and Glover, 2009; Gunbin et al., 2011).

Despite the virtually pan-eukaryotic conservation of the core machinery, the cell cycle regulatory program is under selective pressure to fit the requirements of the organism, for which it adjusts and deploys very differently (and often convergently) across different lifestyles. Unicellular organisms rely heavily in the environmental conditions to undergo cell division as a mean of reproduction, with special mentions to photosynthetic organisms, in which the cell cycle is tightly linked to circadian rhythms, or parasitic life forms with intricate developmental cycles. In multicellular organisms like animals and plants, the cell cycle program is executed as a way to form the organism's body plan and structures, to replenish their cell pool, and/or to regenerate damaged parts. To this end, the cell cycle is bound to the requirements of the whole entity, with only certain cell types able to proliferate in a certain way at certain moments of their ontogeny. Ultimately, failures in this regulation are the reason behind cancers, with fatal consequences for the homeostasis of the organism. Additional examples lay in the so many different ways to

undergo cell division, or to couple division to replication or cytokinesis (see *S. pombe*, *P. polycephalum*, and apicomplexans, discussed above).

Regarding the core cyclin-CDK system, several phylogenetic studies have shown that the last eukaryotic common ancestor (LECA) likely had at least one CDK and multiple cyclin families, including members of A- and B- type cyclins, a situation similar to currently extant eukaryotes such as yeasts (Fig. 10) (Cao et al., 2014; Gunbin et al., 2011; Liu and Kipreos, 2000; Medina et al., 2016). However, our confidence about the phylogenetic relationships of cyclins and CDKs is limited. First, Cyclins and CDKs involved at the same cell cycle phases across different organisms can be orthologs (descending from the same sequence in the ancestor, such as Cyclin B, Clb) or homologs in function (such as apicomplexans CRKs and animal CDKs, or D-type cyclins in plants and animals). This is illustrated by the weak phylogenetic signal of some of the phylogenetic trees of these proteins, where most of the conserved residues belong to the PSTAIRE or the T-loop regions in CDKS and cyclins, respectively. Secondly, the numerous bona-fide cases known for CDKs and cyclins mediating other non-cell cycle-related processes in model organisms (such as regulation of gene expression) (Lim and Kaldis, 2013) suggests this trend also applies throughout eukaryotes. Experimental approaches to identify the interactions and timing of expression of these genes throughout the cell cycle could help to determine the identity of cyclins and CDKs. Unfortunately, these studies are uncommon for organisms outside the groups of landplants, yeasts, and human cells (as shown above), despite our increasing knowledge in the cyclin-CDK gene content of different groups of eukaryotes.

Although the number of species where the cell cycle has been studied has slowly increased over the course of years, the periodic transcriptional program has been mostly characterized in only a handful of model organisms, namely budding and fission yeasts (Orlando et al., 2008; Pramila et al., 2006; Rustici et al., 2004; Spellman et al., 1998), plant cells in suspension cultures (Breyne et al., 2002; Menges et al., 2003; Menges and Murray, 2002) and several *Homo sapiens* cell lines (Bar-Joseph et al., 2008; Cho et al., 2001; Dominguez et al., 2016; Grant et al., 2013; Peña-Diaz et al., 2013; Whitfield et al., 2002), a panel weakly representative of the currently known eukaryotic diversity. Although periodic transcriptional programs are found in many different organisms and cell types, the genes affected by cell-cycle-regulated transcription are divergent between distantly related species, likely due to the fact that transcriptional regulation adapts to the environment and lifestyle of each particular cell type or organism. For example, a number

of periodically expressed genes in *S. cerevisiae* peak at the time of budding (Spellman et al., 1998), a phenomenon that does not occur in animal cells. Transcriptional programs show disparity across different cell types within the same species, as previously exposed for human cells (Giotti et al., 2017; Grant et al., 2013). It is unclear to which extent is periodic regulation even relevant for cell cycle progression on each species (Orlando et al., 2008). Many of the periodic genes in these organisms are therefore expressed periodically for reasons that do not apply to the rest, which leads to little conservation in the periodic transcriptional program of the cell cycle (Jensen et al., 2006).

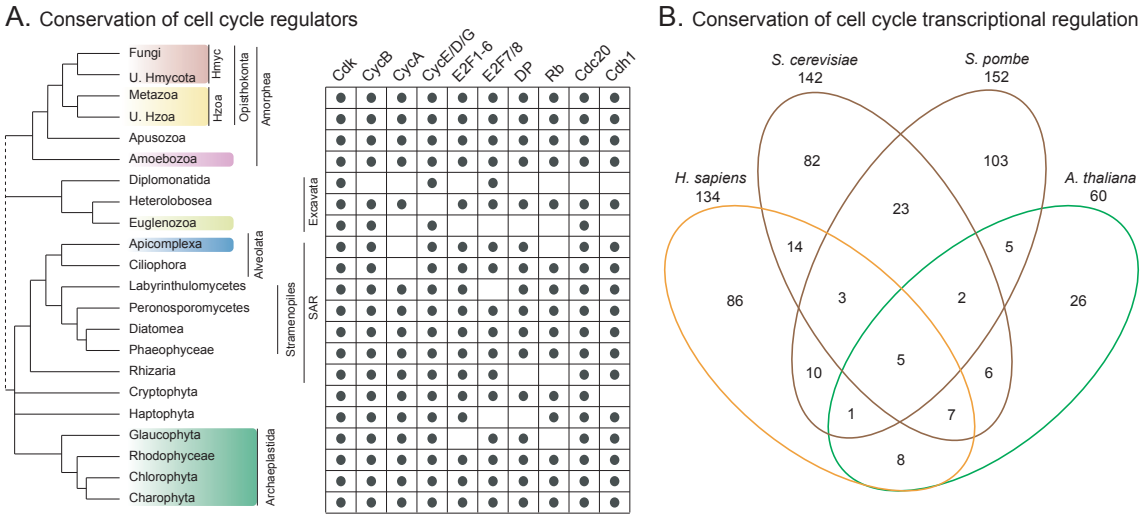


Figure 10: Conservation and evolution of cell cycle regulation in eukaryotes. A: a phylogeny of the major groups of eukaryotes showing the presence/absence of different key cell cycle regulators as described in model systems such as yeasts, human cells, or *Arabidopsis*. 'U. Hmycota' stands for unicellular holomycota, the unicellular relatives of fungi, and 'U. Hzoa' stands for unicellular Holozoa, the unicellular relatives of animals. B: Venn diagram depicting the number of groups of ortholog genes, inferred computationally, that are shared across the periodic transcriptional program of different eukaryote species. Adapted from (Jensen et al., 2006; Medina et al., 2016).

a) The relevance of cell cycle in animal multicellularity

Evolution and implementation of the cell cycle regulation proves important in the emergence and control of new levels of organization, as shown in experimental models of multicellularity (Hanschen et al., 2016; Herron, 2016; Herron et al., 2019, 2018; Ratcliff et al., 2015, 2013). However, cell cycle regulation is so embedded in the life history of species that insights from the current, limited set of organisms fail at providing a comprehensive scenario of how cell cycle regulation evolves in the context of major evolutionary transitions, such as that of the origin of animals.

Some aspects of the cell cycle core system in animals, such as the expansion and subfunctionalization of multiple CDKs, has been suggested to contribute to the robust regulation underlying metazoan cell cycle regulation in a multicellular context (Cao et al., 2014). It is not fully clear what are the evolutionary origins of this regulation, as we are currently missing knowledge from many species within and outside Metazoa. This is reflected in the largely incomplete spectrum of animal diversity: although genomics has allowed to access the gene content of countless animal species, most of the experimental effort has been put into into functional studies of the cell cycle in classic systems such as sea urchin, *Drosophila*, zebrafish, *Xenopus*, and roundworms (See examples and reviews in (Evans et al., 1983; Garcia et al., 2007; Kipreos and van den Heuvel, 2019; Philpott and Yew, 2008; Shepard et al., 2004; Sugiyama et al., 2009)). While the contributions in these studies have provided invaluable information about the elaborate control of the cell cycle in multicellular organisms, all of these species are bilaterians and, on top of it, members of much more diverse bilaterian phyla, highlighting a lack of data from other representatives within their groups such as more protostomes outside Ecdysozoa, and most notably from early-branching metazoan lineages such as sponges, ctenophores, or cnidarians (Laumer et al., 2019).

Logically, most of the work in animal cell cycle has a focus on embryogenic development, under which cell cycle regulation varies greatly among animal embryos (Fig. 5, see above). Our current knowledge comes from experiments often related to other aspects of the cell cycle such as protein biochemistry or developmental patterns; when available, time-series gene expression analyses are rarely exploratory and the sampling resolution does not go down to the level of cell division, only covering the whole organism or tissue of study (Boeck et al., 2016; Bonke et al., 2013; Duffy et al., 2005). Therefore, comparisons of cell cycle regulation are limited to the genetic content and one-to-one

comparisons of gene expression at certain times of development, but no in-depth insights about how cell cycle regulation differs between different animal cell types throughout ontogeny.

Animals belong to the supergroup Opisthokonta, which is formed by Holozoa (animals and their unicellular relatives) and Holomycota (Fungi and their unicellular relatives) (Torruella et al., 2015) (Fig. 1A, Fig. 10A). Within Opisthokonta, the closest unicellular species to animals where cell cycle regulation has been studied are yeasts, which are known to have reinvented a considerable number of otherwise conserved cell cycle regulators upon gene losses, lateral transfers, and other events in their evolutionary history (see above). Because no other unicellular lineage within Opisthokonta has been studied, we lack a clear understanding of the evolution of the cell cycle regulatory program along opisthokonts.

b) Choosing more organisms to study the cell cycle

The evolution of cell cycle regulation faces the classic scientific problem of describing broad generalities (nomothenics) from a pool of unique particulars (idiographics) (Jenner and Wills, 2007). This issue occurs repeatedly across science fields, including other comparative disciplines like evolutionary developmental biology, or evo-devo. Despite this problem, evo-devo studies the fundamental principles whose changes explain the evolution of new body features is possible by cross-comparison of a broad panel of species to extract general tendencies. Making our research of cell cycle regulation extensive to other eukaryotic species might, in the same line, provide a more accurate vision of the evolution of this phenomenon arguably pivotal in the transition to multicellularity. Expanding our research to other lineages requires their establishment as new model systems where to study the cell cycle.

Model systems, or model organisms, are species extensively studied to understand a biological process, hoping that conclusions extracted are also applicable to other organisms (Goldstein and King, 2016; Jenner and Wills, 2007; Leonelli and Ankeny, 2013; Müller and Grossniklaus, 2010; Russell et al., 2017). Historically, life sciences has relied in model organisms for the discovery and description of many crucial concepts, from chromosome-based inheritance and development with *Drosophila* to the understanding of cell biology and the cell cycle with yeasts (Hartwell et al., 1974; Morgan et al., 1915). The choice of a species as a model organism is related to several characteristics that make research easier and transferrable. Model organisms are usually

easy and cheap to grow, maintain and transport in a systematic manner. Ideally they are able to grow axenically under culture conditions, meaning no other species, either known or unknown, grow in the culture or are needed by the organism to grow. Model organisms have quick generation spans and life cycles, which facilitate the completion of many experiments in a given time. Experimental replicates are also facilitated by the generation of large, viable offsprings, such as in *Drosophila* or zebrafish reproduction (Ingham, 1997; Nüsslein-volhard and Wieschaus, 1980). Model organisms feature experimental tractability and amenability, which is crucial for their power as genetic tools. They have high mutation rates or are highly susceptible to genetic techniques, either mutagenesis and classical crossings, or other reverse genetics such as transgenesis or genome editing. These are often complemented by a systematic ability to induce different responses by diverse treatments, either chemical (like drugs), physical (temperature, tensions), or environmental (starvation). In the current times where the cost of genomic sequencing has dramatically dropped, an equally essential requirement is the availability of well curated, annotated genomes that serve to characterize the number of genes, intron content, gene sintenies, and rate of evolution among others, all prior to any experimental approach.

Although often serendipitously, the life history of a model organism suits the question under investigation. Some contributing aspects are the phylogenetic position of the species, the retainment of ancestral characteristics, and the presence a particular biological trait relevant to the question. Three examples are the presence of thousands of chromosomes in the ciliate Tetrahymena, which facilitated the study of telomeres (Ahmed et al., 1998); the arrangement of meiotic products in the ascomycete Neurospora, which facilitates the analysis of genetic recombination (Horowitz et al., 1945; Lindegren, 1932); or the ability to form multicellular structures in social amoebae to study the processes of cell communication and differentiation (Leonelli and Ankeny, 2013; Robertson and Grutsch, 1981).

Perhaps most importantly, research in model organisms proves valuable only if supported and maintained by a scientific community that follows upon the work previously done by others. Such a community is fostered by the establishment of infrastructures of communication and exchange that serve both as the channel for dissemination and as the entity retaining the value of the organism. Examples are found in databases such as YeastDB, WormBase, FlyBase, The *Arabidopsis* Information Resource (TAIR), or more generic ones such as ENSEMBL or Uniprot (Bateman et al.,

2017; Leonelli and Ankeny, 2012). The relevance of the community only contributes to defining it as an essential requirement. Concomitantly, the value of a model organism is also measured by the possibility to transfer the knowledge and skills generated to other model organisms, in order to facilitate comparative approaches.

In addition of the aforementioned aspects, a suitable organism to study cell cycle regulation needs to be experimentally tractable under at least the three following conditions:

1. First, because cell cycle events occur over time in a sequential manner within a single cell, the organism must allow experimentation and monitoring at the cellular level, and with a relatively high degree of temporal resolution, meaning that time-series experiments below the time of generation are essential to assess cell cycle dynamics over time.

2. Second, because cells in a population follow their internal rhythms regardless of their neighbouring cells as a consequence of internal fluctuations, each of them is often at a different stage of the cell cycle. Methods to measure events occurring at the cellular level must be developed; because single-cell approaches were not always possible, synchronization of the cell cycle in cultures has been classically used to address cellular events in a temporal order. Cell cycle synchronization methods include the usage of mutant strains, condition-induction such as starvation or harsh temperatures, or chemical methods such as inhibitors of the cell cycle. Inhibitors are molecules that prevent progression of the cell cycle by disrupting the regulation of the cell, often interfering with cytoskeletal components, biosynthesis enzymes, or replication/division machineries. They are classified as irreversible (if their action permanently prevents progression of the cell cycle upon removal of the agent) or reversible (if a transient effect is produced as long as the agent is present). Reversible inhibitors of the cell cycle are essential for in vivo approaches.

3. Third, because the amount of DNA within a cell changes over time and proves an useful indicator of cell cycle progression, DNA content in the cells of the organism must be accessible and measurable over time, in order to identify the stage of progression of cells during the cell cycle.

Ideally, methods of measurement and data collection shall include cell imaging to track down the different events of mitosis, gene expression analysis in order to assess the periodic transcriptional program, protein interactions in order to elucidate their relationships, protein localization, and mutant/overexpression phenotypes.

Establishing model organisms to wonder about the nature of cell cycle regulation outside animals and yeasts requires us to turn our attention to other lineages within Opisthokonta; if possible, to other organisms more closely related to metazoans. Fortunately, we live in a time where the phylogenetic relationships between animals, fungi, and their unicellular relatives have long been clarified (King, 2004; King et al., 2008; Ruiz-Trillo et al., 2008, 2004; Shalchian-Tabrizi et al., 2008). Multiple lineages are now recognized as the closest living unicellular relatives of animals, and group together in the clade Holozoa.

1.6 The unicellular relatives of animals: where to study the cell cycle?

Holozoa is the group formed by animals and their extant unicellular relatives (CAVALIER-SMITH, 1987; Hehenberger et al., 2017; Lang et al., 2002; Ruiz-Trillo et al., 2008, 2004; Torruella et al., 2015, 2012). The first organisms proposed as unicellular ancestors of animals were choanoflagellates, based on their morphological resemblance with a sponge cell type called choanocyte (Clark, 1866). Advances in genomics uncovered the lineages Filasterea and Ichthyosporea as closely related to animals and choanoflagellates (Shalchian-Tabrizi et al., 2008) (Fig. 11). As mentioned earlier, genomic studies showed that these species contained many genes and protein domains previously thought to be exclusive to animals, thus dating their origin prior to the origin of metazoan multicellularity. Studying these organisms can provide insights into the roles of these genes in a unicellular system, much like those of the unicellular ancestors of animals.

a) Choanoflagellates

Choanoflagellates form a group of more than 250 of spherical heterotrophic, predatory flagellates ubiquitously found as free-living across aquatic environments. They feature a collar of microvilli surrounding an apical flagellum, used for both locomotion and feeding (Leadbeater, 2015). There are two currently available genomes, those of *Monosiga brevicollis* (King et al., 2008) and *Salpingoeca rosetta* (Fairclough et al., 2013). The latter has been extensively characterized at the transcriptional level, and more transcriptomic data exists from many species of choanoflagellates (Richter et al., 2018). Genomic analyses have shown that *S. rosetta* contains two genes from the CDK1 family, and *M. brevicollis* only contains one (Cao et al., 2014).

Choanoflagellates show unique life cycles with asexual and sexual reproduction (Levin and King, 2013), the formation of extracellular siliceous structures, and most interestingly the ability to form colonial structures of multiple cells remain together after cell division. An example are the rosette-like structures of S. rosetta, formed in the presence of a bacterial sulfonolipid (Alegado et al., 2012; Woznica et al., 2016). These findings hint at a deep connection between switching from a unicellular to a multicellular cell/life cycle and environmental cues.

b) Ichthyosporea and Corallochytrea

Ichthyosporea and Corallochytrea group together in the clade Teretosporea (Grau-Bové et al., 2017). Ichthyosporea form a group of more than 40 species often found as saprophytic/osmotrophic parasites, with occasional reports of free-living species (Glockling et al., 2013; Mendoza et al., 2002). In Corallochytrea, the only species described is named *Corallochytrium limacisporum* (Raghu-kumar, 1987). Currently available genomes of these groups include those of *Creolimax fragrantissima* (de Mendoza et al., 2015; Marshall et al., 2008), *Sphaeroforma arctica* (Dudin et al., 2019; Jøstensen et al., 2002), and *C. limacisporum* (Grau-Bové et al., 2017; Torruella et al., 2015).

Ichthyosporeans are frequently multinucleated, and together with *C. limacisporum* follow a broadly conserved developmental program whereby nuclei divide synchronously as the cell increases in volume. After the formation of a coenocyte, nuclei travel to the periphery and cellularization takes place, after which the mother cell bursts and daughter cells are spread as propagules. Occasionally, daughter cells appear in the form of amoebas that disperse and form a new colony (de Mendoza et al., 2015; Dudin et al., 2019; Mendoza et al., 2002; Ondracka et al., 2018). The life cycle of ichthyosporeans shows many resemblances to the coenocytes and syncitia found in some animal embryos, in the slime mould *P. polycephalum* (see above)(Fig. 5), and some fungi (Bonner, 2009; Chen et al., 2007; Grosberg and Strathmann, 2007), providing an interesting system where to study this developmental program.

c) Filasterea

Filasterea is the sister group to animals and choanoflagellates. With only four currently described species (Hehenberger et al., 2017; Owczarzak et al., 1980; Ruiz-Trillo et al., 2004; Stibbs et al., 1979; Tong, 1997), it is, so far, the least rich and abundant holozoan lineage (del Campo et al., 2015; del Campo and Ruiz-Trillo, 2013). Filastereans have typically been described as small, spheric, free-living amoebae bearing multiple, branching filopodia, with occasional flagellum, able to feed by both predation and osmotrophy (Adl et al., 2019, 2012). The most extensively studied species within this group is *Capsaspora owczarzaki*. C. *owczarzaki* (from here onwards referred to as *Capsaspora*) is a filopodiated amoeba originally isolated from the insides of the freshwater snail *Biomphalaria glabrata*, during the study of the parasitic trematode

Schistosoma mansoni within its snail host (Owczarzak et al., 1980; Stibbs et al., 1979). *Capsaspora* amoebae, yet unidentified, were reported to ingest the parasite's sporocysts (Stibbs et al., 1979). Years after deposited in the the American Type Culture Collection (ATCC® 30864) (Owczarzak et al., 1980), it was identified as a close unicellular relative of animals (Hertel et al., 2002); later confirmed by protein coding genes (Ruiz-Trillo et al., 2008, 2004). Its genome was annotated shortly after (Suga et al., 2013), and went under extensive research in the following years (reviewed in (Ferrer-Bonet and Ruiz-Trillo, 2017; Sebé-Pedrós et al., 2017)). Filastereans have been described to display aggregate formation, a mechanism similar to the amoebozoan *D. discoideum* or other sorocarpic amoebae, such as the holomycota *Fonticula alba* (Brown and Silberman, 2013; Robertson and Grutsch, 1981; Worley et al., 1979).

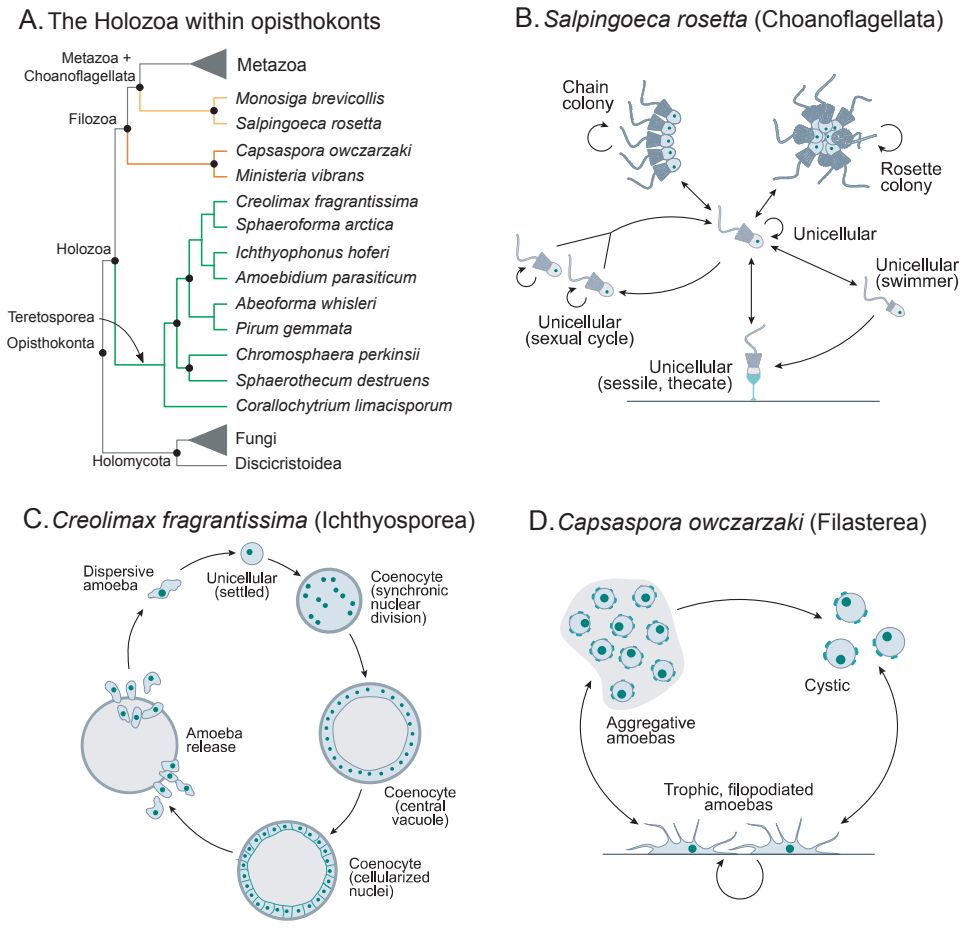


Figure 11: The extant unicellular relatives of animals. A: a phylogeny of the different subgroups of unicellular relatives of Metazoa: choanoflagellates (yellow), filastereans (orange), and teretosporea (ichthyosporeans and *Corallochytrium limacisporum*; green). B-D: The currently described life cycles of the choanoflagellate S. rosetta, Creolimax fragrantissima, and *Capsaspora owczarzaki*.

1.7 Emerging systems in Holozoa: *Capsaspora*

Overall, the three currently described lineages of unicellular holozoa are of great interest for the scientific community to understand the origin of animals, as proven by many genomic and functional studies shedding light into the field (see (Brunet and King, 2017; Sebé-Pedrós et al., 2017) for detailed reviews). While generating knowledge in multiple lineages can complement the findings of each other and help to reconstruct the nature of the unicellular ancestor, their development as model systems is crucial to deepening in the question. Below I present a broad description of findings in *Capsaspora*, and their implications to consider this species as a candidate for an emerging model system.

Capsaspora is a filopodiated amoeba of 6-7µm in diameter that is able to grow in axenic cultures. Under such conditions, its life cycle begins with cells thriving while attached to the bottom of the plate by using their filopodia, in a stage called 'adherent' or 'filopodial'. Cells divide at a ratio of once every approximately twelve hours, increasing the population size until reaching confluency. When resources are depleted, *Capsaspora* cells shrink in size, retract their filopodia and adopt a resistance form called 'cystic' stage, which is able to endure over harsh conditions such as lack of nutrients or desiccation. Perhaps most interestingly, *Capsaspora* adherent cells are able to actively form multicellular structures by aggregation of independent cells. In this stage, cells synthesize a substance similar to an extracellular matrix that has been proposed to contribute to maintain the aggregate structure (Sebé-Pedrós et al., 2013).

These stages have been described at the transcriptional, proteomic, and epigenetic level: in all cases, each stage has a clearly different profile of gene expression, protein abundance and phosphorylation, and associated changes in the chromatin state that are correlated to each other (Sebé-Pedrós et al., 2016a, 2016b, 2013). In adherent cells, genes related to cascade signalling -such as tyrosin-kinases and G-protein receptors- are upregulated, together with transcriptional regulators such as the bZIP superfamily (Sebé-Pedrós et al., 2013). Cystic cells show upregulation of ubiquitin-signaling pathways, vesicle transport, and autophagy, pointing to large rearrangement inside the cell. In aggregates, many genes related to multicellular behavior in animals are upregulated, such as the integrin adhesome, laminin, fibronectin, and associated tyrosin kinases, together with the tubulin cytoskeleton (Sebé-Pedrós et al., 2013, 2010). These different expression programs are accompanied by differential expression and regulation of transcription factors, which may orchestrate the overall differences between the

programs. In adherent cells, the bZIP superfamily is upregulated; in cystic, bHLH superfamily; and in aggregates, some transcription factors linked to metazoan immunity, development, and proliferative signals, such as RUNX and NFκB (Sebé-Pedrós et al., 2013). Protein abundance and phosphorylation also changes across the life cycle of *Capsaspora*, correlating to the transcriptomic data (Sebé-Pedrós et al., 2016b) and are argued as a distinctive hallmark of Holozoa at the onset of the genomic expansion of the tyrosin-kinase families (Suga et al., 2012). Concomitant to this, *Capsaspora* cells have differential regions of open chromatin across stages, which correlate to the presence and placement of different transcription factors and their differential regulation across life stages (Sebé-Pedrós et al., 2016a).

Interestingly, numerous findings in *Capsaspora* shows traits ideal to study the cell cycle. *Capsaspora* seems to conserve most of the currently known, pan-eukaryotic cell cycle regulators: cyclins, CDKs, APC components, Aurora kinases and others, have all been annotated and show a great degree of sequence conservation. This includes cell cycle regulators lost in fungi, such as E2F and Rb (Suga et al., 2013). *Capsaspora* cyclins have broadly been characterized in subfamilies A, B, together with other subfamilies tangentially or non-related to cell cycle regulation (Cao et al., 2014; Medina et al., 2016). The latter applies also to CDKs, with incongruencies in the previous literature about the presence of a cell cycle CDK homolog (Cao et al., 2014; Eme et al., 2011; Medina et al., 2016; Suga et al., 2013). Regarding molecular findings, it is in adherent cells where we find upregulation of cell cycle genes, both at the level of gene expression and protein abundance (Sebé-Pedrós et al., 2016b, 2013). Examples include genes and proteins related to cyclin binding, DNA replication, and G1/S transition. When addressed for DNA content, *Capsaspora* adherent cells under proliferative conditions show a classic example of a bimodal distribution: one subpopulation of cells in G1 coexist with another of G2/M cells, with a small remaining fraction S phase cells (Sebé-Pedrós et al., 2013) (Fig. 9D). In contrast, cells at the cystic stage show a drastic decrease of cells at S and G2/M phases, indicating no proliferation is occurring at that time. Additionally, that same study showed that Hydroxyurea was able to inhibit cell proliferation of *Capsaspora* cells (Fig. 9E). Hydroxyurea inhibits cell cycle progression by acting as a competitor substrate for the enzyme ribonucleotide reductase, necessary to increase the dNTP pool upon entry into DNA replication (Koç et al., 2004; Slater, 1973). Cells are stalled at the G1/S transition in a reversible manner, as removal of the compound allows to continue into the cell cycle (Slater, 1973) (Fig. 9B).

Due to its phylogenetic position, its genomic content, its easiness for culture and experimental tractability, and the extensive knowledge of its life cycle at the molecular level, *Capsaspora* stands as a strong candidate for a model system in the group of unicellular relatives of animals. Together with the availability to address its progression through the cell cycle, to generate gene expression data throughout its life cycle, and with initial evidences of susceptibility to cell cycle inhibitors, proves promising to be used as a model organism to study cell cycle regulation in another lineage outside animals and yeasts.

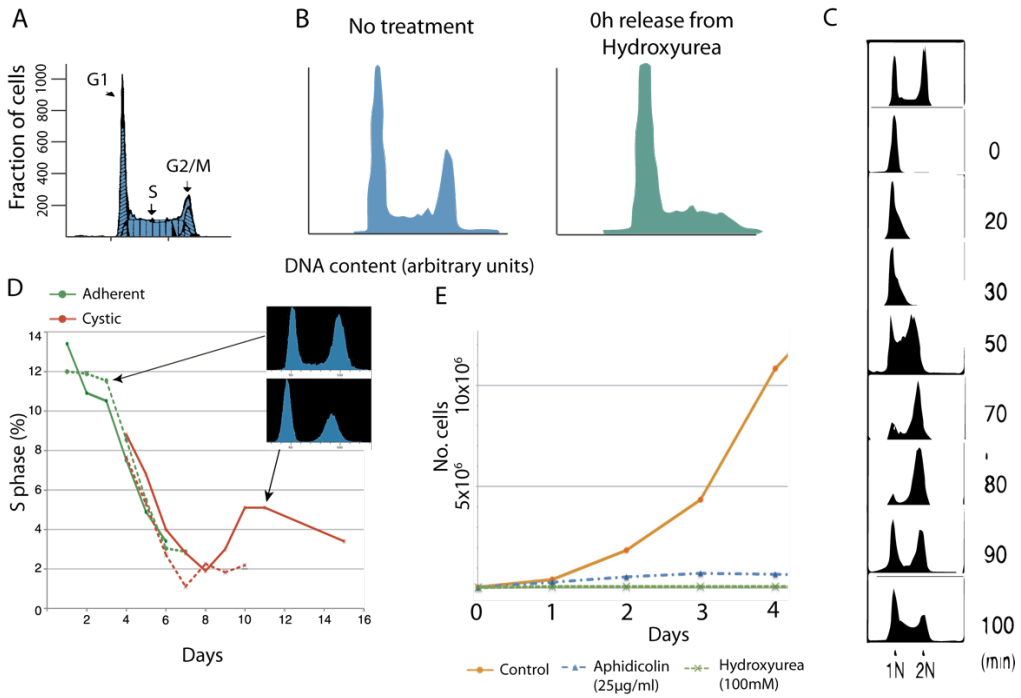


Figure 12: *Capsaspora* and its experimental tractability for cell cycle studies. A: A flow cytometry histogram of an unsynchronized population of human cells. Each peak represents a fraction of cells at a different stage of the cell cycle. B: Effect of hydroxyurea in unsynchronized cultures: following A, the peak of cells corresponding to G2/M cells drops in a culture treated with hydroxyurea, as these cells progress onto the following cell cycle stages while pre-replicative cells remain stuck at the G1/S transition, due to the replicative stress caused as an effect of hydroxyurea. C: Hydroxyurea is a reversible inhibitor of the cell cycle: A culture of *S. cerevisiae* cells was synchronized using hydroxyurea, and followed up after release from this inhibitor. Flow cytometry profiles show that cells progress throughout the different cell cycle phases, completing a whole cell cycle. D: Culture growth of *Capsaspora* measured by flow cytometry and DNA content. E: effect of hydroxyurea on the proliferation of *Capsaspora* cells. A was adapted from (Pozarowski and Darzynkiewicz, 2005); B was adapted from (Borel et al., 2002); C was adapted from (Wilsker et al., 2008); D and E were adapted from (Sebé-Pedrós et al., 2013).

a) Development of genetic tools in *Capsaspora*

Because genetic approaches are not yet possible in *Capsaspora*, a second remaining challenge is the establishment of any sort of genetic tool that would allow experimental approaches at the DNA level. Genetic tools can be used to induce (overexpress), downregulate (knock-down), or inhibit (knock-out or gene silencing) the expression of one or more genes (Bosher and Labouesse, 2000; Prelich, 2012; Ran et al., 2013). Among the most widely used tools lays mutagenesis and mutant isolation through genetic crosses. Mutagenesis and crosses have been classically used in fungi, flies, and other systems throughout the past century in order to unravel the function of every gene based on their phenotype. This technique is made possible when the species shows both a susceptibility to treatment-induced mutagenesis and it has a well-known, systematized sexual cycle. Sexual reproduction has been proved to occur in choanoflagellates and used in the aforementioned approach of genotyping (Levin et al., 2014); however, in the case of *Capsaspora*, no sexual life cycle has yet been identified.

In case the first is not possible or sufficient, another technique consists in the introduction of foreign genetic material into the cell, a process called transfection (when referring to eukaryotes). Transfection is inspired by the discovery of a naturally-occurring process where the DNA acts as a transforming factor (hence the name) in fungi and bacteria. Transfection is a transgenesis tool: it has the additional advantage of allowing the introduction of DNA molecules from foreigner species, or the creation of DNA molecules not existing in nature, through the technology of recombinant DNA. This allows for a much more fine-tuned alteration of the organism's phenotype, as well as additional applications like the in vivo monitoring. Transfection can occur through chemical, physical, or biological factors; the two classically used methods are calcium-phosphate precipitation and electroporation. With the advent of genome editing tools, newer technologies are arising such as nucleofection and lentivirus-mediated delivery. Transfection can be stable if the foreign molecule thrives in the forthcoming generations of the organism, or transient if degraded or lost upon dilution by cell division. Due to the lack of classical genetics in *Capsaspora*, and because transgenesis would allow all its unique applications, the latter method results in the optimal path where to focus research efforts.

Original protocols in the decades of 1970 and 1980 took advantage of the formation of DNA precipitates when incubated with calcium and phosphate solutions, resulting in

calcium-phosphate-DNA complexes (Graham and van der Eb, 1973). These particles can be assimilated by cells relatively easily, and later transferred into the nucleus using shock- inducing protocols (such as osmotic shocks with glycerol) (Loyter et al., 1982). Despite the appearance of a variety of techniques in the following decades, such as electroporation, virus-mediated delivery or lipofection, some authors still use the precipitate-mediated method as it is a much cleaner protocol (Jordan et al., 1996; Sun et al., 2013), which works with materials and compounds that are natural to life (calcium and phosphate have important roles taking part in numerous processes, essentially in signaling) and is therefore more tolerable by cells, regardless of their type or species. These reasons prompt the use of shock-inducing methods as a first option when testing nucleic acids delivery protocols in novel cell lineages or unicellular organisms, as it can be learnt from the work in *D. discoideum* (Crowley et al., 1985; Fey et al., 2007; Gaudet et al., 2007; Martens et al., 2002; Pang et al., 1999).

***

Due to the potential of *Capsaspora owczarzaki* to tackle very interesting questions regarding the evolution of the cell cycle and the origin of animals, in this PhD I focused in the development of methods to experimentally address this species at the cell cycle and transgenesis level, subsequently illustrated by our findings about the evolution of cell cycle regulation in the lineage of Opisthokonta.

Remarks and summary of the introduction

1. The study of the origin of animals as multicellular entities involves the study of the emergence and functioning of those traits that make them multicellular. An elaborate integration of the cell division program under a multicellular context is crucial to explain the origin of clonal multicellularity.

2. To that end, studying the cell cycle in unicellular and multicellular species proves important to understand its evolution in the multicellular lineages such as animals.

3. Unfortunately, our current perspective of the cell cycle regulation is incomplete and biased towards the vast knowledge generated in a handful of model species. This panel, although diverse, is not representative of the currently known diversity of eukaryotes; additionally, it shows several flaws, such as the divergences in the cell cycle machinery and regulation arguably expected in distantly related organisms, which make difficult cross-species comparisons of the cell cycle.

4. Expanding our study of the cell cycle towards other species within the lineage of opisthokonts, together with their establishment as model organisms, remains as the next logical step in this line of research.

5. Among alternative opisthokonts where to study the cell cycle, unicellular holozoans stand out due to their key phylogenetic position to understand the origin of animals, and their genomic content full of genes and features previously thought to be animal-specific. These groups are choanoflagellates, ichthyosporeans, and filastereans.

6. Among filastereans, *Capsaspora owczarzaki* stands out as an emerging model organism due to its experimental amenability and the large knowledge of the molecular basis of its life cycle.

7. From this knowledge, numerous evidences suit this species as a putative candidate where to study cell cycle regulation.

8. The lack of genetic tools in *Capsaspora* precludes any functional study, for which a transfection protocol needs to be developed.

2. Objectives

Objectives

For the reasons exposed in the section 'Introduction', in this thesis I focus in the development of experimental tools and cell cycle knowledge in a unicellular species closely related to animals, the filasterean *Capsaspora*. Building on previous work, the following objectives were pursued:

1. To develop a protocol to establish synchronous cultures of *Capsaspora*.
2. To describe the periodic transcriptional program of *Capsaspora owczarzaki* during its cell cycle.
3. To unravel the cell cycle regulation of cyclins and CDKs in *Capsaspora owczarzaki* throughout its cell cycle.
4. To compare the periodic transcriptional program of *Capsaspora owczarzaki* to those of other species and cell types.
5. To develop a reliable and robust protocol to transiently transfect the filasterean *Capsaspora owczarzaki*.

3. Results

Results

Impact and authorship of the publications

Among the three manuscripts here presented, Alberto Perez-Posada has been the sole first author in Result 3.1, second author in Result 3.2, and co-first author in Result 3.3. The specific contributions of Alberto Perez-Posada to each publication are:

Preprint (section 3.1)

Perez-Posada A, Dudin O, Ocaña-Pallarès E, Ruiz-Trillo I, Ondracka A: *Gradual evolution of cell cycle regulation by cyclin-dependent kinases during the transition to animal multicellularity*. bioRxiv 719534, 2019. doi: 10.1101/719534

- Impact Factor (2019): NA
- 5 Year Impact Factor: NA
- Authorship: first author
- Contribution: Alberto had a leading role in the design of the experiments, the realization of the experiments, the analysis and the interpreation of the data generated, and in the writing and editing of the manuscript and the figures.

Publication (section 3.2)

Parra-Acero H, Ros-Rocher N**, Perez-Posada A**, Kożyczkowska A, Sánchez-Pons N, Nakata A, Suga H, R. Najle S, Ruiz-Trillo I. *Transfection of Capsaspora owczarzaki, a close unicellular relative of animals* Development. 2018. 145(10): dev.162107. doi: 10.1242/dev.162107

- Impact Factor (2016): 5.763
- 5 Year Impact Factor: 6.1
- Authorship: second author

- Contribution: Alberto had a major role in the realization of the experiments, made contributions in the development of the technique, and provided feedback in the writing and editing of the manuscript.

Unpublished review *(section 3.2)*

Ros-Rocher N**, Perez-Posada A**, Ruiz-Trillo I. *The origin of animal multicellularity.* Unpublished

- Impact Factor (2019): NA
- 5 Year Impact Factor: NA
- Authorship: co-first authorship with equal contributions with NRR.
- Contribution: Alberto had a leading role in the writing and editing of the manuscript.

3.1 Result 3.1

Gradual evolution of cell cycle regulation by cyclin-dependent kinases during the transition to animal multicellularity

Perez-Posada A, Dudin O, Ocaña-Pallarès E, Ruiz-Trillo I, Ondracka A. Gradual evolution of cell cycle regulation by cyclin-dependent kinases during the transition to animal multicellularity. bioRxiv. bioRxiv; 2019. p. 719534. DOI: 10.1101/719534

3.2 Result 3.2

Transfection of *Capsaspora owczarzaki*, a close unicellular relative of animals

Parra-Acero H, Ros-Rocher N, Perez-Posada A, Kożyczkowska A, Sánchez-Pons N, Nakata A, et al. Transfection of Capsaspora owczarzaki, a close unicellular relative of animals. Dev. 2018 May 1;145(10). DOI: 10.1242/dev.162107

3.3 Result 3.3

The origins of animal multicellularity

Ros-Rocher N, **Perez-Posada A**, Leger M, Ruiz-Trillo I. *The origin of animal multicellularity.* Unpublished

The origins of animal multicellularity

Núria Ros-Rocher[1], **Alberto Perez-Posada**[1], Michelle Leger[1], Iñaki Ruiz-Trillo[1,2,3]

1 Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Pg Marítim de la Barceloneta, 37-49, 08003 Barcelona, Catalonia, Spain

2 Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Av. Diagonal, 645, 08028 Barcelona, Catalonia, Spain

3 ICREA, Passeig Lluís Companys 23, 08010, Barcelona, Catalonia, Spain

Abstract

The origin of animals involved a transition from a unicellular lifestyle to a multicellular, coordinated organization of cells. Understanding the origin of animals requires to investigate key events in their evolutionary history, such as the emergence of spatial co-occurrence of cells and the origin of processes related to multicellularity, like cell adhesion or communication. Therefore, we need to reconstruct how were the last single-celled ancestor of animals and the first animals, for which we compare early-branching animals and the closest unicellular relatives of animals. In this review, we revise all the recent advances in the characterization of these ancestors with an updated view at the genomic content, complemented by recent advances at the functional level, such as novel insights into the diversity of cell types in early-branching animals and possibly unicellular relatives of animals. These findings prompt us to revise and comment on the latest theories about the transition to multicellularity, which was likely gradual and involved the use of gene regulatory mechanisms in the emergence of early developmental or morphogenetic plans.

1. The origin of animal foundations

Evolution operates by chance and selection, which means there is neither an established blueprint nor a reason to expect continued increases in complexity (Szathmáry and Smith, 1995). Living organisms have thus evolved in many different directions during their time on Earth. There have been important events that seem to have taken place once in the history of life, such as the emergence of eukaryotes; yet, multicellular life forms have appeared repeatedly in an independent manner in at least 16 eukaryotic lineages (Bonner, 1998; King, 2004; Knoll, 2011; Rokas, 2008).

Multicellularity can evolve via two different mechanisms: by cells coming together (aggregation) or by cells remaining attached after they divide (clonality). In aggregative multicellularity, cells from the same species that may or may not have the same genetic background assemble together forming a multicellular structure of a heterogeneous population of cells, often towards the aim of reproduction and dispersion (Du et al., 2015; Schaap, 2011). While aggregative multicellularity has evolved repeatedly across different eukaryotic lineages (Brown et al., 2012, 2009; Brown and Silberman, 2013; Lasek-Nesselquist and Katz, 2001; Worley et al., 1979), clonal multicellularity has appeared on fewer occasions and is responsible for the greatest radiations of life forms

in the tree of life: land plants, fungi, and animals. The latter group shows an unprecedented degree of disparity in forms and behaviours, and the highest levels of organismal complexity, body plan diversity, and developmental regulation (Grosberg and Strathmann, 2007). Nevertheless, the underlying genetic and mechanistic basis which triggered and shaped the evolution of animals remains a long-standing question. To better understand the transition to animal multicellularity and thus their origin from a unicellular ancestor, we need to know more about the nature of the organisms from which animals first evolved and diversified.

In this review, we focus on the reconstruction of the evolutionary stages that were crucial during the transition to animal multicellularity: the last common ancestor (LCA) of animals and the single-celled ancestor(s) of animals (Fig. 1A). We summarize the current knowledge on the genetic toolkit, cell type diversity and ecological context of these stages inferred by comparative genomic analyses between the earliest branching animals and those radiating later, and between animals with their closest unicellular relatives. We finally discuss an updated perspective on the evolution of animals, stressing the gradual complexification and the emergence of complex regulatory mechanisms to drive early developmental programs.

1.1. Reconstruction of the last common ancestor of animals: the fossil record and molecular clocks

All animals living today share a common multicellular ancestor, that we denominate the animal LCA, which followed the evolution of multicellularity from a single-celled ancestor (Fig. 1). This transition occurred more than 600 million years ago (MYA), although there is still no exact consensus between fossil record estimates and biochemical and molecular clock-based estimates. Until recently, fossil record inferences were the only way to address when the animal LCA originated and what it looked like. The first unequivocal evidence of animals in the fossil record date back to the Ediacaran period, with molecular clocks extending the estimate of the emergence of animals 100 million years more, during the Cryogenian and Tonian periods (Antcliffe et al., 2014; Budd and Jensen, 2017; dos Reis et al., 2015; Douzery et al., 2004; Hedges et al., 2004; Knoll, 2011; Knoll and Hewitt, 2013; Narbonne, 2005; Peterson et al., 2004). Some examples of these animal fossils are Dickinsonia, Kimberella (both from around 560 MYA), Eoandromeda octobrachiata (555 MYA) or Eocyathispongia qiania (600 MYA), whose body plans suggest that axial patterning and symmetry were

featured early in animal evolution. Although the use of sterols as biomarkers of animal life in the fossil record remains controversial (Nettersheim et al., 2019), molecular characterization of the sterol composition of sponge-like and *Dickinsonia* fossils suggests that these organisms share ancestry with animals living today (Bobrovskiy et al., 2018; Love et al., 2009). Thus, all extant animal lineages share a common ancestor that lived on Earth prior to the Cambrian explosion, ca. 700 MYA, which later radiated in the Cambrian period (Budd and Jensen, 2017; Narbonne, 2005) (Fig. 1B).

Despite the limited information provided by the fossil record about the biology of the first animals, we can do partial reconstructions by examining the extant, earliest branching animals. All animals have a body plan relying on symmetry at least during a certain stage, which allows a clear distinction between bilaterian (bilaterally symmetrical) and non-bilaterian animals. Under this scenario, and before confirmation by molecular phylogeny, animals with more rudimentary body plans (i.e., poriferans, ctenophores, cnidarians, and placozoans) were considered the earliest branching lineages (Haeckel, 1874). Studies in sponges (Porifera) prompted authors like Muller to speculate about the nature of the first animals (Müller et al., 2004).

1.2. Knowledge of animal genomes helps to reconstruct the animal LCA genetic toolkit

Advances in comparative genomics have provided deeper insight into the molecular characteristics of these lineages with dozens of non-bilaterian transcriptomes and genomes currently available (Chapman et al., 2010; Moroz et al., 2014; Putnam et al., 2007; Riesgo et al., 2014; Ryan et al., 2013; Srivastava et al., 2010, 2008), thus uncovering their genetic toolkit and clarifying their evolutionary relationships. Understanding the evolutionary relationships of metazoan (animal) early-branching lineages illuminates how metazoan traits arose and evolved during evolution (Dohrmann and Wörheide, 2017). Although all animal phyla are well defined both morphologically and phylogenetically, there is no consensus about many of the internal relationships between some of them (Telford et al., 2015), including early-branching animals. The debate over the earliest split in the metazoan tree of life is long and still ongoing, with two main hypotheses under consideration – either Ctenophora or Porifera are the sister group of the rest of Metazoa (Laumer et al., 2019; Moroz et al., 2014; Pisani et al., 2015; Ryan et al., 2013; Telford et al., 2015; Whelan et al., 2015). Both lineages diverged very early in animal evolution and feature unique characteristics

that make them very derived organisms. Previous studies have claimed the full support of each hypothesis, only for a later study to reach conflicting conclusions due to differences in their genes of choice, the sequence positions used in the alignments, the outgroups chosen, and the assumptions of homogeneous or heterogeneous substitution rates (Laumer et al., 2019; Moroz et al., 2014; Pisani et al., 2015; Ryan et al., 2013; Telford et al., 2015; Whelan et al., 2015). Future approaches require the addition of more well curated genomes to improve phylogenomic resolution (Gold et al., 2019; King and Rokas, 2017; Laumer et al., 2019; Leclère et al., 2019). Despite the obscure phylogenetic relationships at the base of the animal tree, for which we remain agnostic, reconstruction of the animal LCA is still possible to a certain extent (King and Rokas, 2017) and our points in this review apply regardless of the scenario considered.

The deep knowledge provided by genome sequencing of several animal species has contributed to reconstructing the metazoan (animal) LCA gene toolkit. Two recent, independent studies addressed the genomic features of the animal LCA using a wide range of currently available genomes (Paps and Holland, 2018; Richter et al., 2018), and concluded that around two thousand gene families originated on the animal stem lineage (Richter et al., 2018). The rate of gene innovation in or immediately prior to the animal LCA is larger than at other key points of the animal stem, highlighting the importance of gene innovation in the origin of animals (Paps and Holland, 2018). Surprisingly, not all the genes detected to have originated in the animal LCA are still present in all extant phyla. In fact, only around 2% of these gene families are conserved across all animals; among early-branching, currently sequenced animals, the sea anemone *Nematostella vectensis* seems to retain most of the animal LCA genetic toolkit (Putnam et al., 2007; Richter et al., 2018). Some of the genes that emerged at the onset of animals take part in processes that differentiate them from other lineages: the set of a few, highly retained genes in all phyla are enriched in functions of DNA binding, Transcription Factors (TFs) (already hinted in previous studies, see (de Mendoza et al., 2013)), innate immunity, and members of the Wnt and TGF-β signalling pathways, the latter linked to the emergence of animals since they determine axial patterning (Adamska et al., 2007; Lee et al., 2006; Leininger et al., 2014; Paps and Holland, 2018; Richter et al., 2018). These families are absent or underrepresented only in highly derived early-branching animals (Chang et al., 2015; Schenkelaars et al., 2017). Besides gene innovation, pre-existing gene families expanded and new domain architectures were acquired in the emergence of animals, *e.g.,* in receptor tyrosine-kinases, SNARE receptors, and homeobox and helix-loop-

helix TFs (de Mendoza et al., 2013; Paps and Holland, 2018; Srivastava et al., 2010), and gradually expanded and increased in the stem towards bilateria (Larroux et al., 2008). It has been argued that expansion and subfunctionalization may have played a role in contributing to increasing complexity in animals (de Mendoza et al., 2013; Larroux et al., 2007, 2006; Srivastava et al., 2010). However, gene loss also appears to have impacted the evolution of animals, with as much gene loss as gene gain in animals compared to their unicellular relatives, affecting pathways such as amino acid biosynthesis and osmosensing (Richter et al., 2018). Comparisons have also been made between animals and other multicellular life forms, although these are complicated by the long evolutionary distances separating them (Nedelcu, 2019). Based on these comparisons, it has been argued that evolution and complexification of animal forms could have been boosted by the acquisition of the SAND domain at the onset of this lineage through lateral gene transfer; this shared domain of animals, plants and multicellular volvocine green algae would explain why all these lineages have an exuberant diversity of cell types and developmental patterns (Nedelcu, 2019).

1.3. Comparative genomics unravels traits of the animal LCA: adhesion, contractility, cell communication and development

Perhaps unsurprisingly, the common gene repertoire of all animals points that cell adhesion, motility, and developmental and nervous signalling were mechanisms already present in the animal LCA. Cadherins (molecules mediating cell-cell interactions) and integrins (mediating cell-extracellular matrix interactions) are found in most of non-bilaterians (Moroz et al., 2014; Putnam et al., 2007; Ryan et al., 2013; Srivastava et al., 2010) (Srivastava et al., 2010b; Ryan, Moroz, Putnam 2007), and a recent study determined the presence of basal lamina components laminin and collagen (or its truncated sister *spongin*) in comb jellies (Ctenophora) and at least one group of sponges (Fidler et al., 2017). As for adherens junction components and cell polarity, sponges have a full, fairly conserved repertoire (Fahey and Degnan, 2010; Nichols et al., 2012; Srivastava et al., 2010) but ctenophores lack Crumbs complex and Scribble homologs (Belahbib et al., 2018). These results altogether point to a rich repertoire of adhesion genes in the animal LCA. Although sponges are sessile during adulthood, two orthologs of heavy chain myosin of striated and non-striated muscle have been identified in two sponge genomes. This suggests that a contractile system existed in the animal LCA, and later evolved into distinct muscle types in different

lineages. This system was potentially lost in sponges and independently gave rise to striated muscle cells in cnidarians, ctenophores, and bilaterians (Burton, 2008; Steinmetz et al., 2012). Similarly, at the organizational level, ctenophores possess an independently-derived mesodermal tissue, despite their lack of key bilaterian mesoderm specification genes (Martindale and Henry, 1999; Moroz et al., 2014; Ryan et al., 2013). This suggests that the regulatory mechanisms necessary for establishing early fates in layers of cells (such as the muscle cells in the ctenophore-specific mesoderm) were present before the emergence of bilaterians. If ctenophores are the earliest branching animals, then these mechanisms would likely have been present in the animal LCA.

Most animals sense and respond to stimuli with complex behaviours thanks to the nervous system. The LIM-associated homeobox family of TFs was originally found to be only in Metazoa (Putnam et al., 2007), and together with genes such as bHLH, Sox and elav, they are among the first genes found in animal genomes related to neural development and neural cell fate. Surprisingly, LIM-associated homeobox genes are present and expressed in developing *Amphimedon* embryos, despite *Amphimedon* lacking any overt nervous system (Srivastava et al., 2010). This same species possesses genes related to postsynaptic densities, as well as different receptors and biosynthesis enzymes of different molecules used as transmitters, the latter linked to the production of secondary metabolites (Riesgo et al., 2014). This pan-representation of neural genes is however limited in some lineages such as ctenophores, where the main messengers used in synapses in other animals are absent (Moroz et al., 2014). This has led some authors hypothesize a parallel evolution of the nervous system in early-branching animals such as ctenophores (Moroz and Kohn, 2016, 2015).

Nervous systems rely on the transmission of information at two main levels: within the same cell (electrochemical impulse) and between cells (chemical or electric synapsis). There are numerous examples in nature of electrically excitable epithelia (Josephson, 2004; Mackie, 2004; ROBERTS, 1969) which suggest a potential common origin of the electrochemical basis observed in nerve impulses and that of epithelia. Neuronal features heavily rely on current transmission derived from osmotic changes (Bucher and Anderson, 2015), and so their evolution required a machinery for osmotic control (i.e., ion channels to generate action potential and a basal lamina) in a continuum of cells delimiting compartments (the outside and the inside of the cell). At the cell-cell communication level, some observations indicate that early-branching animals use

their nervous system to communicate information about their microbiomes and the surrounding bacteria (Klimovich and Bosch, 2018; Musser et al., 2019), sharing a common origin of the foundations of the neural and the immune systems. There is abundant evidence of innate immunity components occurring in different animal lineages, from Toll-like and Ig receptors to TFs and complement system in sponges and cnidarians (Brennan et al., 2017; Gauthier et al., 2010; Miller et al., 2007; Riesgo et al., 2014), some being expressed by cells in intimate contact with neurons.

A particularly important group of signalling pathways are those mediating animal development. Embryonic development coordinates the placement of the aforementioned components into a coherent, physically integrated organizational plan. Several components of the Wnt signalling pathway, involved in anteroposterior axial patterning, are expressed in sponge larvae, during cnidarian development, and in several structures of both adult sponges and adult ctenophores (Adamska et al., 2007; Hobmayer et al., 2000; Lee et al., 2006; Riesgo et al., 2014; Srivastava et al., 2010, 2010; Windsor Reid et al., 2018). Components of other signalling pathways, such as Hedgehog, Notch, or TGF-β, are present but more scattered between lineages and species (Moroz et al., 2014; Riesgo et al., 2014; Ryan et al., 2013). Despite being patchy and incomplete in different species, all animal signalling pathways (including those responsible for patterning bilaterians) are present in virtually all early lineages (Nichols et al., 2006; Paps and Holland, 2018; Riesgo et al., 2014), suggesting the animal LCA contained a machinery with the potential to drive axial patterning in a developmental program.

1.4. The last common ancestor of animals likely possessed many different cell types

A major feature of animals is the spatial distribution of labor between coexisting cells within a single multicellular entity. The building blocks of multicellular bodies are different kinds of cell types (Box 1), each specialized in different roles within the whole organism. Cell types have their own cell modules used in different processes (e.g., contraction, secretion, signalling and reception), that are normally regulated by a well-defined genetic program (a set of TFs and other specific regulatory mechanisms). This implies that some genes are expressed by certain cell types but not others – in other words, they express a limited number of genes encoded in the genome. On the other hand, the totality of genes in the genome of a unicellular organism may be used by a

single cell throughout its life cycle. The genome partitioning in animal cell types reflects an increase of regulation to determine cell fate genetic programs between different cell types within a single organism (Arendt et al., 2019). Understanding the origin of animals therefore requires asking questions about the evolutionary origins of cell types and their mechanisms of differentiation.

All the information provided by comparative genomics is relative to the presence or absence of certain genes in the organism. Therefore, sequencing of the genome alone fails to provide information on how those genes are expressed and regulated in different cells, and on the putative identities of the cells within an organism. Insights in this field have come hand in hand with advances in single-cell genomics. It has been reported that cell type diversity increases within animals in each step of body plan complexification, with relatively low numbers cell type diversity in early-branching animals contrasting the hundreds or thousands of cell types in bilaterians and in vertebrates, respectively. However, new research is showing that cell type diversity is more complex and varied in animals than we thought. Many of the *bona fide* cell types of non-bilaterians have been classically described at the morphological level, but the access at the molecular level has allowed the definition of numerous identities previously unnoticed (Marioni and Arendt, 2017). Identifying previously undescribed cell types in early-branching metazoans can help us understand the role of gene expression regulation in the definition of cell types. Cell atlases of sponges and cnidarians (Musser et al., 2019; Sebé-Pedrós et al., 2018b; Siebert et al., 2019) have allowed the identification of transcriptional signatures in the form of correlated and differential gene expression across the transcriptomes of different cells, leading to the definition of numerous “metacells” in these organisms (Baran et al., 2018; Sebé-Pedrós et al., 2018b). Based on their expression profiles, these metacells can be linked to putative functions, such as muscle or cnidocyte cells in *Hydra* and *Nematostella,* or archaeocytes in sponges (Musser et al., 2019; Sebé-Pedrós et al., 2018b; Siebert et al., 2019; Sogabe et al., 2019), or can be used to detect novel, undiscovered cell types, such as nitric oxide-sensitive contractile cells or neuroid-amoeboid cells in *Spongilla* (Musser et al., 2019). Metacells appear to change across developmental stages of these organisms, with differences in the metacell composition in adult and larval stages of *Nematostella* (Sebé-Pedrós et al., 2018b; Siebert et al., 2019). Likewise, they can be linked to different cell fate trajectories branching from stem cells, reconstructing the cell type phylogeny of the organism (Siebert et al., 2019). One main recognisable signature of these cell types is the expression of TFs: hundreds of them are differentially

expressed in *Nematostella*, of these hundreds 85% are of metazoan origin and half are specific to certain metacells. These numbers reflect the relevance of TFs diversification in the lineage that led to animals (de Mendoza et al., 2013). Some of these TFs coexpress with genes whose promoters contain binding sites matching them (Siebert et al., 2019). Correlation of gene expression between TFs suggests there exists a hierarchy of control to define the identity of the cells, meaning the control of cell type identity emerges from the control at the tissue level, from a broader to a more specific fate. Distal regulation may also play a role in the specification of cell types, as many of the cell type specific genes have associated putative enhancers spanning multiple kilobases upstream (Sebé-Pedrós et al., 2018b).

An especially large number of cell types correspond to neuronal identity in these species (Sebé-Pedrós et al., 2018b). These cells show extensive conservation in neuronal effector gene repertoires with other animal species, but not at the level of coexpression, supporting the existence of an ancestral cell type with neuronal characteristics, at least at the onset of cnidarians and bilaterians. Interestingly, a group of *Spongilla* cells named *neuroids* surround other cells and express a combination of TFs and effector genes related to immunity and synaptic communication (Musser et al., 2019). Animals like *Nematostella*, *Hydra* and *Spongilla* have thus been able to use a basal regulatory machinery and different ancestral cell modules to evolve their own types of cells. This fact reinforces the possibility of an animal LCA machinery related to that of nerve and immune cells which could have later diverged independently in each lineage, giving rise to similar structures, such as the nervous system.

So far, single-cell approaches are bringing a much deeper insight into cell type differentiation in many early-branching Metazoa, by adding sponges, comb jellies and placozoans to the comparisons (Sebé-Pedrós et al., 2018a). Such analyses have revealed how sponge larvae and adult organisms exhibit different cell types, an unexpectedly high diversity of cell types in ctenophores, and the presence of peptidergic cell types in placozoans (Sebé-Pedrós et al., 2018a). Interestingly, these species show differences in the control of cell type-specific gene expression, such as a large diversity of TFs and very specific promoter motifs, versus the use of distal elements (Sebé-Pedrós et al., 2018a). Because of these differences in very different lineages of early branching animals, it is likely that cell type genome regulation in the animal LCA involved a heterogeneous program of distal regulation and complex TFs networks. However, the large evolutionary distances and high divergence between

these lineages to the ancestor makes it difficult to find homologous cell types between them. Only very conserved aspects, such as the expression profile of epithelial or ciliated cells, are conserved across species. Most of the genes expressed in all cell types in these species are either found in all eukaryotes or are lineage-specific, and in all cases it appears that novel genes are more limited to specific cell types (Sebé-Pedrós et al., 2018a).

From this scenario of poor conservation and high divergence we can make three observations: first, the very limited living record of early branching animals becomes evident. The fossil record shows that many animal groups present in the Cambrian explosion went extinct leaving no representatives to this day, a situation with little prospect of change due to the discovery of only three animal phyla in the past thirty years (Kristensen, 2002), and with all the documented metazoan diversity by environmental metabarcoding belonging to already existing phyla (López-Escardó et al., 2018). Second, the currently extant early-branching animals are very derived and specialized, suggesting that secondary losses and simplification fostered the evolution of new structures and genes in these lineages ((O'Malley et al., 2016), see Box 2). Third, while tracing the origin of specific cell types to the origin of Metazoa proves a difficult task, these approaches show that animals have been able to diversify and specify the fates of their cells in many different ways but using the same mechanisms of genomic regulation and the same effector machineries responsible for cell physiology. We can conclude that the animal LCA was already rich in cell types which share some of their cellular foundations with those in extant species.

1.5. Reconstruction of other ancestors of animals

It is important to note that the animal LCA and the first common ancestor of animals may have not been the same (Box 1). In fact, the first animal, defined as a multicellular entity with a developmental plan, likely gave rise to other lineages that subsequently went extinct prior to the divergence of all modern animal lineages from the animal LCA. Research is so far limited to the reconstruction of the animal LCA, and the different unicellular ancestors of animals.

However, we can have an approximate reconstruction of the first common ancestor of animals based on our current knowledge of the animal LCA. For instance, taking into account the patchy distribution of all those machineries in all the different lineages of

early-branching animals, we can infer that the genetic toolkit of the animal LCA was very rich in genes related to metazoan innovations, from the cellular foundations necessary to generate epithelia-like layers to neuron-like signalling cells and muscle-like contractile cells. Although specific cell types may have evolved convergently in different early-branching lineages (Moroz, 2015; Steinmetz et al., 2012), all of their genomes encode elements of the genes and domains that allowed these cell types to emerge. Many animal-specific pathways and mechanisms were thus largely complete in the animal LCA (similar to the observations about the cnidarian-bilaterian LCA by (Putnam et al., 2007)), suggesting that those machineries were likely present in previous ancestral states possibly even in the first animals (Fig. 1, Box 1). Similarly, based on our inferences of cell type diversity in the animal LCA, those ancestors prior to the animal LCA likely had the ability to regulate cell differentiation by means of hierarchical TF networks and distal regulation in different cells within the multicellular collective, which translates to a certain degree of spatial cell differentiation. Rather than a drastic bloom of innovations, it is likely that gene expansion, co-option, increased regulatory sophistication and a transition from temporal to spatial gene regulation had a crucial impact on the gradually increasing complexity of animals ((Sebé-Pedrós et al., 2017), and references within).

Still, phylogenomic studies and analysis for the reconstruction of the animal ancestors are limited by the data currently available for such comparisons: genomic data of early-branching animals is limited to a handful of species, which may or may not be good representatives due to events of gene loss and fast evolution. Likewise, our findings would be biased towards the assumption of numerous innovations in the animal lineage unless we include other lineages in our comparisons. For these reasons, the study of the origin and evolution of animals requires sequencing more early-branching animal genomes and, just as importantly, to expand our focus to other lineages outside Metazoa.

2. A unicellular perspective into animal origins

In the last decade, the advent of molecular and sequencing techniques and genomic and morphological characterization of unicellular species closely related to animals has drawn a robust phylogenetic framework of animals and their unicellular relatives. Animals are closely related to a heterogeneous assembly of unicellular lineages, together comprising the Holozoa clade within the eukaryotic group Opisthokonta

(Cavalier-Smith, 1987; Lang et al., 2002; Ruiz-Trillo et al., 2008, 2004; Shalchian-Tabrizi et al., 2008; Torruella et al., 2015, 2012) (Fig. 1). By comparing their biology and genome content to those of animals, we can reconstruct the cellular foundations of animal evolution and infer the minimal genomic complexity of common ancestors of animals. Moreover, unicellular relatives of animals, hereafter unicellular holozoans, allow a more comprehensive reconstruction of the gene content, regulatory mechanisms and potential lifestyle(s) of the single-celled ancestor of animals, which are key to understanding the transition from unicellular to multicellular forms.

2.1. Unicellular holozoans, animals' closest relatives

The closest unicellular lineage to animals comprises a group of more than 250 species of spherical/ovoid heterotrophic predatory flagellates known as Choanoflagellatea, originally associated to animals more than a century ago by their morphological resemblance to *choanocytes*, a specific cell type of sponges (Clark, 1866). This similarity, together with their phylogenomic confirmation as the closest living unicellular relatives of animals, gave rise to hypotheses of animals evolving from a choanoflagellate-like ancestor (Cavalier-Smith, 2017; Cavalier-Smith and Chao, 2003; Lang et al., 2002; Ruiz-Trillo et al., 2008; Steenkamp et al., 2006; Wainright et al., 1993; Zettler et al., 2001). Choanoflagellates are widely distributed worldwide in a range of aquatic environments and are considered to be major contributors to aquatic microbial food webs (Arndt et al., 2000; del Campo and Massana, 2011; del Campo and Ruiz-Trillo, 2013; King, 2005; Nitsche et al., 2007; Thomsen et al., 2016; Torruella et al., 2015; Wylezich et al., 2012). They present an apical funnel-shaped feeding collar that forms a complex of interconnected actin-filled microvilli at the base of a single flagellum, which allows generating water currents to capture and phagocytose bacterial prey (Leadbeater, 2015). Some of their representatives are remarkable for producing extracellular siliceous basket-like coverings (Carr et al., 2008; Leadbeater, 2015) and others are able to form multicellular-like structures of stably adherent cells, known as *rosettes,* as a product of oriented cell divisions from a single founder cell (Dayel et al., 2011; Fairclough et al., 2010). Another interesting behaviour has been observed in the recently described *Choaneca flexa,* which is able to form enormous cup-shaped colonies (Brunet et al., 2019). These colonies reversibly invert their curvature in response to light through a rhodopsin-cGMP pathway, representing a similar behaviour to concerted movement and morphogenesis in animals and the first observation of light-response in a choanoflagellate (Brunet et al., 2019).

Filasterea is the second closely related lineage to animals, composed of only four known species to date: the endosymbiont *Capsaspora owczarzaki* (Owczarzak et al., 1980; Stibbs et al., 1979), the marine free-living heterotroph *Ministeria vibrans* (Cavalier-Smith and Chao, 2003; Kiss et al., 2009; Mylnikov et al., 2019; Shalchian-Tabrizi et al., 2008; Tong et al., 1998; Torruella et al., 2015) and the more recently described freshwater predatory flagellates *Pigoraptor vietnamica* and *Pigoraptor chileana* (Hehenberger et al., 2017). Similarly to choanoflagellates, some filastereans are able to form multicellular structures. For example, the filopodiated amoeba *C. owczarzaki* presents a multicellular stage formed through the active aggregation of independent cells, representing the first example of an aggregative behaviour among unicellular holozoans (Sebé-Pedrós et al., 2013b). *C. owczarzaki* also presents two additional stages which, together with the aggregative stage, are all differentially regulated at the transcriptomic, proteomic and phosphoproteomic levels (Sebé-Pedrós et al., 2016a, 2016b, 2013a; Stibbs et al., 1979). Recently, clusters of few cells have also been detected in *M. vibrans* (Mylnikov et al., 2019) and in both *Pigoraptor* species (Hehenberger et al., 2017).

The next lineage is Ichthyosporea, a group of osmotrophic and saprotrophic fungus-like protists frequently multinucleated and sometimes with a single posterior flagellum, especially in dispersal forms. Most members of Ichthyosporea have been found in commensal, mutualistic or parasitic relationships with aquatic (both freshwater and marine) and terrestrial animals, and therefore most of them have been directly isolated from different animal tissues, especially guts of molluscs and arthropods (Glockling et al., 2013). Few species have been identified as free-living (Hassett et al., 2015). Over the past decade, the number of taxa in this group has increased considerably, reaching to date about 40 characterized species, although around half of them are phylotypes (Feldman et al., 2005; Lohr et al., 2010; Lord et al., 2012; Marshall et al., 2008; Marshall and Berbee, 2011). Moreover, many taxa in the Eccrinales and Amoebidiales originally identified as fungi have been phylogenetically classified as ichthyosporeans (Benny and O'Donnell, 2000; Cafaro, 2005; Mendoza et al., 2002). Some species exhibit fungal-like traits (Mendoza et al., 2002; Torruella et al., 2015) or other complex phenotypes, such as motile pseudopodia, hyphal or plasmodial structures (Marshall and Berbee, 2011). Interestingly, some members of the Ichthyosporea present a broadly conserved developmental mode consisting of large, multinucleated spherical or ovoid coenocytes, often with a large central vacuole or multiple vacuoles and a thick cell wall. At some point, this stage releases multiple motile spherical or limax-shaped

amoebas by cellularization of the internal nuclei (Dudin et al., 2019; Mendoza et al., 2002; Ondracka et al., 2018). Interestingly, in *Sphaeroforma arctica,* nuclear division operates independently of cell size by a clock-like mechanism, generating a self-organized polarized layer of cells (Dudin et al., 2019; Ondracka et al., 2018).

Finally, the Corallochytrea clade includes a single species, *Corallochytrium limacisporum* (Raghu-kumar, 1987), a small spherical-shaped free-living osmotroph, originally isolated from marine coral reefs in the Arabian Sea and more recently in Hawaii (Raghu-kumar, 1987; Torruella et al., 2015). *C. limacisporum* presents a complex developmental mode that in some aspects resembles the one present in Ichthyosporea. Uninucleated cells undergo a number of binary cell divisions without cytokinesis, causing a single large cell or colony to form and grow until it releases limax-shaped amoebas that disperse and form new colonies (Raghu-kumar, 1987). Although it is still not clear whether *C. limacisporum* is able to undergo a coenocytic cycle, some observations revealed it is able to undergo serial cell divisions with palintomic cleavage (Butterfield, 2011; Chen et al., 2014; Xiao et al., 2012). In the most recent and taxon-rich phylogenomic analysis, it was classified as the sister group to Ichthyosporea, forming a monophyletic group named Teretosporea, referring to the prevalence of “rounded spores” in both *C. limacisporum* and ichthyosporeans (Grau-Bové et al., 2017; Torruella et al., 2015). However, analysis of the newly described species *Syssomonas multiformis* (Hehenberger et al., 2017) grouped them together in an independent clade branching between Filasterea and Ichthyosporea*,* named Pluriformea (Hehenberger et al., 2017). *S. multiformis* is a freshwater-dwelling predatory flagellate that feeds on large eukaryotic prey, especially on heterotrophic chrysomonads and bodonids (Hehenberger et al., 2017). Morphologically, it presents a complex developmental mode that includes amoeboflagellate, amoeboid cells and motile swimming cells. Similar to the filasterean *Pigoraptors*, it can form clusters of multiple cells and spherical cysts. Both *C. limacisporum* and *S. multiformis* phylogenetic position need to be further investigated.

Other clades falling within or related to different unicellular holozoan lineages have been identified using molecular data from environmental surveys, a powerful tool to describe new diversity (Arroyo et al., 2018; del Campo et al., 2016; del Campo and Ruiz-Trillo, 2013; Pawlowski et al., 2016, 2012, 2011). For example, in deep-sea sediments, Operational Taxonomic Units (OTUs) from Choanoflagellatea and

Ichthyosporea were mostly abundant in sediments from the Arctic Sea (Pawlowski et al., 2011). A recent study identified more than 60% of OTUs as potential new taxa at the species level in unicellular holozoans, especially to *S. multiformis* (Arroyo et al., 2018). This study surveyed freshwater environments, which are generally considered to have more ecological niches for eukaryotic communities than marine environments (Arroyo et al., 2018). Filastereans have been recently identified in environmental surveys ((Heger et al., 2018); Arroyo et al., unpublished). Altogether, this points to a wider geographical and ecological distribution of key taxa and potentially new groups related to unicellular holozoans, suggesting that still there is a considerable hidden diversity of unicellular holozoans to be uncovered (Arroyo et al., 2018; del Campo et al., 2015; del Campo and Massana, 2011; del Campo and Ruiz-Trillo, 2013; Pawlowski et al., 2011; Savin et al., 2004; Stougaard et al., 2002; Takishita et al., 2007, 2006, 2005).

2.2. The single-celled ancestral toolkit

Analyses of the genomes, transcriptomes, proteomes and gene regulatory mechanisms of unicellular holozoans make it possible to identify key features that are present/shared between extant unicellular holozoan species and present/absent in the animal lineage. These features can be inferred to be present in the unicellular ancestor and thus allow reconstructing the nature of the single-celled ancestral toolkit.

Currently, there are two available choanoflagellate genomes, those of *Monosiga brevicollis* (King et al., 2008) and the colony-forming *S. rosetta* (Fairclough et al., 2013), which has also been thoroughly characterized at the transcriptomic level and cell biology level (Fairclough et al., 2013; Levin et al., 2014; Levin and King, 2013; Woznica et al., 2017, 2016). Transcriptomes are available from a broad representation of different choanoflagellate species (Brunet et al., 2019; Fairclough et al., 2013; Richter et al., 2018). Among filastereans *C. owczarzaki* has been more deeply studied as the complete genome, transcriptome, proteome, phosphoproteome and epigenome are available (Denbo et al., 2019; Sebé-Pedrós et al., 2016a, 2016b, 2013b; Suga et al., 2013). There is transcriptomic data available for *M. vibrans* (Torruella et al., 2015) and both *Pigoraptor* species (Hehenberger et al., 2017), and analysis of their genomes are underway. To date, there are six available ichthyosporean genomes: *S. arctica* (Grau-Bové et al., 2017); *C. fragrantissima* (de Mendoza et al., 2015); *I. hoferi* (Torruella et al., 2015) and *P. gemmata, A. whisleri* and *C. perkinsii* (Grau-Bové et al.,

2017). Currently, transcriptomic data is available for several species, including *S. destruens* (Torruella et al., 2015) and *C. perkinsii* (Grau-Bové et al., 2017) and *C. fragrantissima* (de Mendoza et al., 2015), *A. whisleri* and *P. gemmata* (Torruella et al., 2015), *A. parasiticum* (Torruella et al., 2012) and *S. arctica* (Ondracka et al., 2018; Torruella et al., 2012). Transcriptomic data is also available for *S. multiformis* (Hehenberger et al., 2017) and the genome of *C. limacisporum* was recently sequenced (Grau-Bové et al., 2017).

Unicellular holozoans possess and express an unexpectedly large repertoire of genes that have key functions in animal multicellularity showing that these genes predate animal origins (de Mendoza et al., 2015; Fairclough et al., 2013; Grau-Bové et al., 2017; Hehenberger et al., 2017; Richter et al., 2018; Sebé-Pedrós et al., 2016a, 2016b, 2013b, 2011, 2010; Suga et al., 2013; Torruella et al., 2015, 2012). For example, almost 400 gene families previously thought to be animal innovations were identified in choanoflagellates (Richter et al., 2018). These observations make gene innovation alone unable to fully explain the transition to animal multicellularity. Thus, in addition to the emergence of novel gene families only found in animals, gene co-option and/or regulatory changes to pre-existing ancestral genes may have been important driving forces for the transition from unicellular to multicellular forms (de Mendoza et al., 2015; Fairclough et al., 2013; Grau-Bové et al., 2017; King et al., 2008, 2008; Richter et al., 2018; Sebé-Pedrós et al., 2011; Suga et al., 2013; Torruella et al., 2015, 2012). Discovering the function of these and other “multicellularity-related” genes in unicellular holozoans and how they are regulated during their life cycles would be the next crucial step towards understanding how those genes were co-opted during the transition (Box 3) (Brunet et al., 2019; de Mendoza et al., 2015; Goldstein and King, 2016; Rokas, 2008; Ruiz-Trillo et al., 2007).

2.2.1. Cell adhesion toolkit

Early discussions in the field pointed out that evolution of cell-cell and cell-extracellular matrix (ECM) adhesion mechanisms was needed prior to the emergence of animals (Müller et al., 2004). Cell adhesion mechanisms are essential for building a cohesive and physically integrated assembly of multiple cells, tissues and organs in animals. Nevertheless, both cell-cell and cell-ECM adhesion molecules are present in unicellular holozoans. For example, choanoflagellates possess several predicted cadherin domain-containing proteins (Abedin and King, 2010; Nichols et al., 2012) and c-type

lectins localised into the interiors of *rosettes* in *S. rosetta* as part of an extracellular matrix, which potentially mediates intercellular adhesion (King et al., 2008; Levin et al., 2014). Strong evidence for several domains present in the animal homolog *Flamingo* have also been identified in several choanoflagellate species as well as several Protocadherin domains identified in five species of choanoflagellates (Richter et al., 2018). Most components of the integrin adhesome system, essential for cell-ECM junctions in animals, have also been detected in unicellular holozoans. An integrin beta domain was identified in the choanoflagellate *D. costata* and the integrin beta-binding protein ICAP-1 was present in all choanoflagellates except *M. brevicollis* and *S. rosetta* (Richter et al., 2018). Interestingly, some *S. rosetta* cells within *rosette* colonies are connected with other neighbouring cells by intercellular bridges (Laundon et al., 2019).

Among filastereans, *C. owczarzaki* possesses and expresses the complete integrin adhesome system (Sebé-Pedrós et al., 2010; Suga et al., 2013) and other ECM-related proteins, which are upregulated in aggregates (Suga and Ruiz-Trillo, 2013). Nearly a complete integrin adhesome system and the associated signalling and cell adhesion components, such as the integrin-linked kinase (ILK), PINCH and Parvin are also present in both *Pigoraptor* species (Hehenberger et al., 2017) and integrin alpha and beta genes are present in *C. limacisporum* (Grau-Bové et al., 2017). Several components of this system are also present in *S. multiformis* (Hehenberger et al., 2017). Other key elements present in animal ECM, such as canonical Type IV collagens, were found in *M. vibrans* (Grau-Bové et al., 2017).

2.2.2. Gene regulatory networks: cell signalling, communication, transcription factors and genomic regulation

As previously mentioned, the architecturally complex body plans present in animals require the coordination and arrangement of cells. The developmental programs responsible for the body plan are deployed through TFs, signalling molecules, and other proteins interacting in gene regulatory networks (GRNs). Cell signalling pathways mediate cell-cell communication and environmental sensing/response. Metazoan intracellular signalling components were present in the unicellular ancestor of animals, but upstream receptors and ligands evolved after the transition (Sebé-Pedrós et al., 2017). For example, key genes involved in spatial signalling include Hedgehog, WNT, TGF-β and JAK-STAT networks, which are absent in unicellular holozoans. Nevertheless, a recent study revealed that Notch, Delta and homologs of the animal

Toll-like receptor genes evolved prior the animal-choanoflagellate divergence as they are present in several choanoflagellate species (Richter et al., 2018). Another group of signalling molecules, the receptor tyrosine kinases (RTKs), might have been key to the evolution of the holozoan lineages provided their evolutionary history in this group. Dozens of RTKs have diversified independently in all groups of unicellular holozoans (Manning et al., 2008; Suga et al., 2014, 2012), while others, such as the cytoplasmic tyrosine-kinases Src, FAK and CSK, also have premetazoan origins but likely evolved later (Schultheiss et al., 2014, 2012; Suga et al., 2012; Suga and Miller, 2018). Other members of important signalling cascades, such as the Myc-Max network (Young et al., 2011) and intracellular components of the Hippo signalling pathway are present in *C. owczarzaki* their metazoan upstream receptors (*i.e.,* Crumbs and Fat) are absent (Sebé-Pedrós et al., 2012).

Transcription factors (TFs) drive gene expression by binding to the DNA at specific recognition sequences, generally localized in the proximity of the transcription start site of the target gene. From the numerous TFs known to be present in the animal LCA, many originated in the previous unicellular ancestors or even in Opisthokonta. LIM-associated Homeobox TFs, formerly thought to be only in animals, are present in several ichthyosporeans, in *Capsaspora* and in *Corallochytrium* (Grau-Bové et al., 2017). Other examples include Nuclear Factor-κB, the p53/63/73 family, RUNX, and T-box TFs (de Mendoza and Sebé-Pedrós, 2019; Sebé-Pedrós et al., 2011). As with some components of the Hippo pathway (Sebé-Pedrós et al., 2012), some of these TFs already display the potential to play a part in gene regulatory networks well established in Metazoa, such as Brachyury and MYC (Sebé-Pedrós et al., 2016a), highlighting that this regulatory role was likely already present in the unicellular ancestors of animals.

In animals, gene expression can be regulated by modifying DNA accessibility. Compaction and relaxation of the chromatin can foster the recruitment of TFs and other regulatory proteins. Many of the genes involved in controlling chromatin accessibility are also present in unicellular holozoans, such as the histone acetyltransferase p300/CBP or many histone post-translational modifiers (Grau-Bové et al., 2017; Sebé-Pedrós et al., 2016a). Interestingly, the life stages of *Capsaspora* have associated changes of chromatin accessibility in the proximal cis-regulatory regions, in contrast to the abundant distal regulation in animals (Sebé-Pedrós et al., 2016a). Overall,

comparative studies have inferred that the unicellular ancestor of holozoa was able to use GRNs and epigenomic mechanisms to regulate the transition between life stages.

2.3. The evolution from the unicellular ancestral genome

Most unicellular holozoans, such as *Corallochytrium, Capsaspora* or *Chromosphaera*, have relatively compact genomes; with a few exceptions in Ichthyosporeans (Grau-Bové et al., 2017), which like animals typically have larger genomes than their unicellular relatives (*e.g.,* early-branching animals genomes are 300-500 Mb) (Elliott and Gregory, 2015; Simakov and Kawashima, 2017). This indicates that the LCA of Holozoa may have had a considerably compact genome, which was progressively shaped by differential retention of ancestral states and secondary innovations. Independent selective pressures on different traits, such as genome size, synteny conservation and reorganization and intron density, were key during the transition to multicellularity. Moreover, there was also a continuous process of protein domain gain and diversification involving multicellularity-related genes in animal ancestors (Grau-Bové et al., 2017). Altogether, two major genome expansions, before and after the origin of animals, had a major impact on gene families expansion and diversification which, together with Transposable Element propagation, may have contributed to increases in the size of animal genomes (Grau-Bové et al., 2017; Paps and Holland, 2018).

Importantly, as outlined above, gene novelty had a considerable impact during the origin and evolution of animals, especially in biological functions key for animal multicellularity, such as genes involved in differential gene regulation (*e.g.,* TFs, signalling pathways), cell adhesion (*e.g.,* cadherins), cell-type specification, cell cycle and immunity (Paps and Holland, 2018; Richter et al., 2018; Richter and King, 2013). Nearly 400 gene families previously thought to be animal-specific instead evolved prior to the animal-choanoflagellate divergence, as illustrated in the earlier section (Richter et al., 2018). Novel gene regulatory mechanisms, such as enhancers and chromatin-structural modifications (distal regulatory elements) were key to the rise of metazoans (Sebé-Pedrós et al., 2016a). An increase of transcriptional regulatory networks also made an important contribution to gene family diversification. Some classes of TFs expanded to give rise to new families at the onset of metazoa, including T-box, SOX, homeobox, and bHLH. An additional level of acquired transcriptomic regulatory complexity, including alternative splicing events by exon shuffling, exon skipping or

intron retention (Grau-Bové et al., 2017; Liu et al., 2005), contributed to novel sources for transcriptomic innovations (Barbosa-Morais et al., 2012; Grau-Bové et al., 2017; Irimia et al., 2009; Nilsen and Graveley, 2010, 2010; Rogozin et al., 2012).

In parallel, a recent study suggests extensive gene loss or divergence occurred during metazoan evolution, especially in genes related to enzymatic functions (*i.e.,* kinases), although we have to consider that taxon sampling was limited (Paps and Holland, 2018). Nevertheless, this points to a higher turnover of genes and greater genomic plasticity during the diversification of animals (O'Malley et al., 2016). These changes in the architecture and content of the genome reshaped the genetic toolkit, structure and regulatory mechanisms of the earliest animals genomes (Grau-Bové et al., 2017; Simakov and Kawashima, 2017). Probably an early burst of gene diversity in the LCA of Holozoa, enriched in TFs and cell adhesion machinery, was followed by multiple and differently-timed episodes of synteny disruption, intron gain and genome expansions. The evolution of non-coding genes and novel epigenetic mechanisms may have also played major roles during the transition to animal multicellularity (Sebé-Pedrós et al., 2016a). For example, the appearance of developmental promoters and distal enhancer elements was potentially a key innovation at the onset of animals, increasing *cis-regulatory* complexity (Sebé-Pedrós et al., 2016a). Moreover, a recent study reported the first evidence for miRNA genes and homologs of the animal miRNA biogenesis machinery (including Drosha and Pasha) in Ichthyosporea (Bråte et al., 2018). This indicates a unicellular origin of animal miRNAs and the associated microprocessor complex, which evolved long before the last common ancestor of animals, early in holozoan evolution (Bråte et al., 2018).

Thus, the unicellular ancestor of animals already possessed several gene regulatory mechanisms, including regulation of chromatin states, complex cis-regulation by enhancers, cell-type-specific alternative splicing and post-transcriptional regulation of mRNA translation via miRNAs with remarkable similarities to those of animals, suggesting that the foundations of animal genome architecture and gene control mechanisms were laid before the origin of multicellularity (Bråte et al., 2018; Grau-Bové et al., 2017; Sebé-Pedrós et al., 2016a). Inclusion of new genomes in the future may change the reconstruction of the ancestral genome and regulatory capacities because comparative genomic studies are highly sensitive to taxonomic biases.

3. Our current perspective on the origin of animals

To define what makes an animal, classical comparisons started comparing animals with unicellular organisms like yeast (Aravind and Subramanian, 1999; Hazkani-Covo et al., 2004). Everything absent in yeast was notated as an important innovation to the origin of animals. The study of numerous early-branching animals and unicellular relatives of animals has gradually changed our perspective, by revealing that many genes and mechanisms thought to be animal innovations are also present in unicellular organisms, and were therefore likely already present in their (unicellular) shared common ancestor with animals. Concomitantly, numerous studies have increased our knowledge about the environment in which animals originated and diversified. The definition of what makes an animal has changed concordantly to these discoveries, as have the hypotheses concerning what were the crucial innovations, or circumstances that led to the origin of animals.

The study of early branching animals has allowed us to reconstruct some of the main characteristics of the last common ancestor of animals. The animal LCA was likely an aquatic, obligate multicellular organism composed of different cell types with cell modules resembling those seen in the main cell types of extant animals. The animal LCA likely featured cell-cell adhesion using cadherin and cell-ECM adhesion using integrin-related proteins, collective movement by cell contractility, environmental sensing and cell-cell communication via synaptic-like modules, and an epithelium-like cell layer composed at least partly of collar cells used to capture bacteria as a food source (Brunet and King, 2017; Richter and King, 2013). It probably reproduced sexually via sperm and egg and presented a form of developmental processes through mechanisms of cell division, cell differentiation and invagination present in all animals (Brunet and King, 2017; Richter and King, 2013). Such diversity of cell types and complex organization was regulated by a diverse set of TFs and epigenomic machineries involving distal regulation, and the initial steps of development likely involved Wnt and TGF-β genes acting coordinately in signalling pathways.

In turn, the study of the unicellular relatives of animals has allowed us to infer that the last common ancestor of Holozoa was likely a bacterivorous organism with a compact genome containing genes involved in cell adhesion and communication, and regulatory mechanisms involving chromatin remodelling, alternative splicing, TF networks, and *cis*-regulation of the genome. The different lifestyles observed among unicellular

holozoans, and especially the evidence of regulated temporal cell differentiation prior to the emergence of animals (de Mendoza et al., 2015; Fairclough et al., 2013; Sebé-Pedrós et al., 2016a, 2013b), indicates that the ancestral unicellular state was relatively plastic, with a complex life cycle integrating transient cell identities, or states, at the temporal level. These included, for example, a differentiated sedentary filter-feeding or heterotrophic life stage and a proliferative stage that could have included dispersive forms. Interestingly, recent and ongoing efforts in the study of multicellular structures in the unicellular relatives of animals are investigating the possibility that cells with different morphologies and identities, driven by different genetic programs, might coexist in these stages ((Laundon et al., 2019); Najle et al., personal communication). This has been also hinted when performing comparative transcriptomics between sponge cell types and unicellular holozoa, with the pluripotential archaeocyte resembling *S. rosetta* colonies or *C. fragrantissima* coenocyte (Sogabe et al., 2019). Therefore, the single-celled ancestor of animals likely possessed the foundations to establish different cell type identities.

Events explaining the transition

Our increasing knowledge of these species illustrates that these two evolutionary stages key for animal evolution had many things in common. The emergence of some aspects of their complex biology, which overall contributed to the emergence of animal multicellularity, can be explained by evolutionary events at the genomic level. Whole genome duplications, gene duplications and gene family expansions, lateral gene transfers, or invasions of TEs increased the complexity of their genome architecture, in turn increasing the evolvability of the genome by distal regulation, neofunctionalization and co-option of genes, and increases in the cross-talking of regulatory networks (possibly throughout dimeric TFs (Sebé-Pedrós et al., 2017)). This may have contributed to a rudimentary integration of temporal cell types into a single physical entity.

Additionally, it has been argued that external factors and ecological triggers were equally as important as internal genomic changes during animal evolution (Paps and Holland, 2018). One example is the biogeochemical context in which animals originated and diversified, with changes in ocean chemistry such as the availability of iron and copper (Canfield et al., 2008, 2007; Knoll and Carroll, 1999; Runnegar, 1982) or the great oxygenation event occurred around 700 MYA, although some authors

argue the latter was not as critical (Lyons et al., 2014; Mills et al., 2014; Mills and Canfield, 2014). As multicellular organisms, the origin of animals can also be related to all the advantages of being multicellular. The emergence of new ecological niches (Stanley, 1973), or selection for multicellularity as an escape from predation, were potentially other driving forces for the origin of animals (Boraas et al., 1998; Herron et al., 2019).

The ecological context might have also had an impact in animal evolution, such as in shaping animal feeding modes and morphological features (Sperling et al., 2013). For example, animals evolved in an environment teeming with bacteria, and have lived since then in close association with bacteria throughout their evolutionary history. In fact, host-associated microbiota can actually regulate development and gut morphogenesis in animals (Klimovich and Bosch, 2018). In this context, being in a close relationship with bacteria could have potentially impacted animal evolution by requiring a system of cell communication to harbour bacterial symbionts and commensals and a defense system to deal with bacterial pathogens. Bacterial interactions are also observed among the closest unicellular relatives of animals, especially among choanoflagellates. *Rosette* development in the emerging model *S. rosetta* has been well characterized, and is known to be triggered and enhanced by a bacterial sulfonolipid (Alegado et al., 2012; Carr et al., 2017, 2008, p. 200; Dayel et al., 2011; Fairclough et al., 2013, 2010; Nitsche et al., 2011; Woznica et al., 2016) and regulated by the *rosetteless* gene (Levin et al., 2014). Bacterial lipids regulate developmental switches both activating or inhibiting *rosette* formation in *S. rosetta* (Woznica et al., 2016). This is not the sole example of environmental bacteria playing a key role during its life stage, as *S. rosetta* is capable of sexually reproducing upon induction by a bacterial chondroitinase (Levin and King, 2013; Woznica et al., 2017, 2016). Interestingly, some of these bacterial species also regulate life cycle transitions in other animal species (Nyholm and McFall-Ngai, 2004). Numerous investigations in other choanoflagellate species highlight the role of bacterial interactions in choanoflagellates (Alegado and King, 2014; Woznica and King, 2018). An example is *Salpingoeca monosierra* which presents 10 bacterial symbionts within a large single colony that reaches more than 100 µm in diameter, the first known choanoflagellate microbiome (Hake et al., unpublished). It is likely that living in an environment of bacteria likely fostered the foundations of holobiomes from the beginning of animals, and that the origin of some essential aspects of animals can be found in the interactions with microorganisms.

Theories

Former definitions of ‘what is an animal’ involved coordinated at the multicellular level, the presence of spatial cell differentiation, and a coordinated developmental plan starting from a single cell. Thus, the theories explaining the origins of animals involved the acquisition of mechanisms to generate epithelium-like multicellular structures. Further studies and comparisons revealed that the mechanisms underpinning these features were likely elaborated in the stem lineage of animals, building upon pathways and features present in their single-celled ancestor (Brunet and King, 2017; Richter and King, 2013). Revised theories proposed the acquisition of spatial regulation as one of the main drivers of the origin of animals, in contraposition to the temporal regulation of cell types exhibited by their unicellular relatives(Mikhailov et al., 2009; Zakhvatkin AA, 1949) Concomitant to the fact that unicellular relatives of animals likely possess different cell types at the multicellular stage, we propose an updated revision of what changes might have been key to the emergence of animals. In our view, multicellular structures, with different cell type identities coexisting, were likely present prior to the origin of animals (Fig. 2).

We envision an initial scenario as shown in Zakhvatkin in 1949 and revised in Mikhailov et al., 2009, of an organism with a complex ontogeny of different life stages transiently regulated in a temporal way. Each stage consisted of different cell types using distinct cell modules to perform different roles: substrate attachment, feeding, swimming, or mating are some examples. One of these stages was a multicellular structure originating through clonal division, displaying spatial coexistence of different, non-committing cell identities driven by unique genetic programs of transdifferentiation. In this stage, different living functions (feeding, motion, and secretion) occurred simultaneously as they were performed by different cells. Whereas spatial regulation itself was formerly proposed as the main driving change, these features may have likely been present in the so-called ‘unicellular ancestors of Metazoa’ and cannot justify in itself the coordinated multicellularity of animals.

Below we speculate about some aspects that may have played a key role in the origin of animals, in relation to some of their features and in no particular order, and always in the context of incremental complexity discussed in this review.

Increased genomic innovation and co-option of pre-existing elements:

The transition was accompanied by an increase of genomic innovation, including many new, rapidly evolving, and subsequently widely conserved genes encoding regulatory functions associated with animal multicellularity: gene regulation, signalling, cell adhesion and cell cycle. Nevertheless, co-option of and regulatory changes to pre-existing elements present among unicellular holozoans set the foundations for further gene family expansions and diversifications. This in turn contributed to an increased layer of regulatory mechanisms for cell-type specification in the animal stem lineage, and likely played a major role in the sections below.

The acquisition of axial patterning

The current unicellular relatives of animals so far identified do not seem to show any particular arrangement of cells when forming a multicellular structure, seemingly lacking coordination with their neighboring cells, unlike animal development. The developmental plan of virtually all animals is driven by morphogens and signalling pathways, with the main components of these pathways mostly absent in the unicellular relatives of animals .

Signalling molecules driving modular responses are present in bacteria and other eukaryotes, and are used by unicellular organisms to respond to quorum (Cámara et al., 2002), to arrange components within the cells (Hachet et al., 2011; Martin and Berthelot-Grosjean, 2009), or to trigger major developmental plans in aggregative organisms (Robertson and Grutsch, 1981). The unicellular ancestor already possessed different receptors, ligands and gene regulatory mechanisms to respond to extracellular cues in a changing environment. One crucial change could have been the co-option or neofunctionalization of these genes to respond to signals originating within the multicellular organism, leading to the ability of cells to integrate spatial information stemming from neighboring cells (Fig. 2). The pathway in question likely was able to trigger a modular non-binary response, which would explain multiple outcomes, and at least one component of at least one mechanism of genome regulation capable of triggering different cell type identities. Likewise, assymmetrical cell division might have been important at the onset of animals to generate spatial differences between cells. A primary axis likely arose as a result of the ability to arrange cell types. One potential

candidate could be the Wnt/β-catenin signalling pathway, known to regulate the anteroposterior axis of the body plan even in early-branching animals, providing a nucleating architecture for the different cell types to arrange (Hobmayer et al., 2000; Windsor Reid et al., 2018). These primary axis may have led to the formation of simple morphogenetic plans (Sebé-Pedrós et al., 2017). Thus, functional modules in the organism would turn out to be based on spatial groups of cells, instead of individual coexisting cells.

Progressive acquisition of cell type identity

The integration of temporally regulated to spatially coexisting cell types could have contributed to a gradual regionalization of functions that in turn fostered the emergence of morphogenetic programs (Sebé-Pedrós et al., 2017). Originally labile, cell type identity (and in turn GRNs) became increasingly dependent of within-group constraints and less dependent on external factors, leading to a certain commitment of cell fate. This might have occurred by the GRNs of the cells becoming more linked or dependent on the signals within the organism, overriding the autonomy of the cell to respond to the environment by transdifferentiation. The emergence of cell types would allow selection to operate at the level of individual cells in terms of the collective fitness, constituting a fine-tuning of the within-group selection by (Szathmáry and Smith, 1995). Inherently, the emergence of multicellular structures might have enhanced the differences between cells in different regions or parts of this multicellular entity.

The emergence of a conjoined gene regulatory program of fertilization and multicellular development

Animals produce very distinct kinds of gametes. Gamete fusion determines initial polarity and triggers the developmental program in animal eggs (Ajduk and Zernicka-Goetz, 2016; Fatema et al., 2019), meaning it could have served in earlier stages of animal evolution as an early device for asymmetric cell division, generation of a rudimentary axis, and establishment of cell fates. During development and other moments of the ontogeny, animal cells are able to proliferate in response to signals from within the organism, by controlling the entry into the cell cycle. The set of cell cycle regulators *Capsaspora* shares some traits with those of animals, with some conserved transcription factors related to proliferation and the timing of key genes from animal checkpoints during their cell cycles (Perez-Posada et al., 2019; Sebé-Pedrós et al., 2016a), although these organisms lack the genes required to trigger cell cycle

progression in response to extracellular signalling in animals (Cao et al., 2014; Medina et al., 2016). So far, we do not know of any unicellular holozoan where the formation of the multicellular stage is linked to the fusion of gametes, and the latter seem to be identical between sexes. At some point during in the stem lineage leading to animals, an ancestor with the ability to both generate a multicellular morphogenetic plan through axial patterning and perform sexual reproduction likely integrated these two programs in a single developmental plan (Fig. 2).

Relegation of unicellular stages in favor of the multicellular stage

In relation to the context discussed above, the multicellular stage could have prevailed over the unicellular stage, likewise fostered by escaping from predators, enhanced resource exploiting, and a relaxation of ecological constraints due to increases in the availability of some nutrients. Likewise, the emerging properties concomitant to multicellularity, such as the division of labor, could have favored the thriving of the multicellular stage as the proliferative stage (Sebé-Pedrós et al., 2017). Unicellular stages could have become simple forms of dispersions, if not only exclusive to the haploid form prior to fertilization, as we observe in today's animal gametes.

Thus, the origins of animals involved a long, gradual evolutionary process rather than a one-step evolutionary jump, which paved the path to develop animal multicellularity by coupling complex development, sperm-egg fusion, and serial cell division throughout the integration of spatial cell differentiation (Brunet and King, 2017; Sebé-Pedrós et al., 2017).

4. Concluding remarks

In the past years, a vast body of knowledge produced by the advent of molecular omics has elucidated the relationships between animals' earliest-branching lineages, and those of animals with their closest unicellular relatives. In this review, we have gone through all the major points of discussion in the current time: from an obscure origin where one cell became many, we have changed to the view of an intricate, complex single-celled ancestor of animals showing many of the traits currently found in animals, potentially even the spatial coexistence of different cell types. The convergence of this field with the molecular characterization of cell types using single cell technologies is proving essential to understanding the origin of animal multicellularity.

In order to expand this knowledge, we must aim at getting to know more about these lineages. On the one hand, phylogenies used to understand the origin of animals might be made more robust by the inclusion of more genomes from porifera, ctenophora, and other early-branching animals. Our knowledge of the gene content of the animal LCA might also benefit from the inclusion of more genomes, and studies at the functional level of individual cell types has proven that there is still a lot to know about the cell type diversity and the hidden biology of early-branching animals. On the other hand, the identification of more species in the branches of unicellular relatives of animals, together with the sequencing of their genomes, can help to uncover the evolutionary history of some animal genes still not well known. Likewise, studying the biology of these organisms might help to uncover additional aspects that could have been important for animal origins. Understanding the multicellular stages of these species involves understanding how they emerge within the life cycle, how they are maintained and organized, what are the different cell types within the collective, what are the temporal and (if any) spatial mechanisms to regulate their different identities. Their life cycles and ecological context might be further investigated by studying the possible relationships of these species with neighboring microorganisms, the temporal transitions between cell types, and the mechanisms of sexual reproduction. Overall, all of these areas of study might help to identify the origin and evolution of any sex-associated developmental program involving a multicellular stage. Finally, the advent of genetic tools in these organisms will foster the study of the specific functions of genes that might participate in these processes, such as transcription factors, adhesion molecules, and components of different signalling pathways and other animal modules (such as cell contractility, synaptic proteins, and developmental pathways).

With this narrowed set of questions, and the onset of a new era of functional and single-cell-based approaches within the framework of the origin of cell types, we find ourselves excited about - but most importantly eager to begin - unravelling the origins of animals.

BOXES

BOX 1: Terminology used in this review

- **Holozoa:** eukaryotic group comprising animals, choanoflagellates, filastereans, and teretosporeans. The largest group including *Homo sapiens* but not *Neurospora crassa* (Adl et al., 2019).
- **Last common ancestor of animals (== animal LCA):** the ancestral stage from which all animal phyla living today radiated. Reconstructed from features present in and shared by all animals. Undoubtedly presenting all the features shared by all animals, including complex, coordinated multicellularity. Therefore, it can be classified as an animal.
- **First multicellular ancestor of all animals (== animal FCA):** Partly reconstructed from features shared between early diverging animal lineages (i.e., sponges, ctenophores, placozoans and cnidarians), even if they are absent in bilaterians (Richter and King, 2013). This ancestor lived subsequent to ALL the changes that led to the foundations of multicellularity in animals, and may or may not be the same as the animal LCA (we cannot know).
- **Single-celled ancestor of animals:** the single-celled ancestor immediately preceding the first common ancestor of animals originated.
- **Urchoanozoa:** the last common ancestor of animals and choanoflagellates. May or may not be the same as the single-celled ancestor of animals.
- **Last common ancestor of Holozoa:** The ancestor shared by Metazoa, Choanoflagellatea, Filasterea, Ichthyosporea, and Chorallochytrea.
- **Metacell:** in single cell genomics, subgroup of homogeneous scRNASeq profiles with only local variance relative to the total dataset, useful for clustering and quantitative gene expression analyses (Baran et al., 2018). Ultimately, it can be related to certain cell types, but only upon experimental validation.
- **Cell type:** in its simplest definition, a cell type was defined as a unit of classification to distinguish forms of cells according to different morphologies or phenotypes. Cell types are often related to different germ layers during the formation of the embryo, with nerve and epithelial cells coming from the ectoderm, muscle and blood cells from the mesoderm, and gut cells from the endoderm (Burton, 2008; Chang and Hemmati-Brivanlou, 1998; Lewis and Tam, 2006). Whereas vertebrate cell types are often defined by their committed fate and being unable to de-differentiate, cells from early-branching animals are known to transdifferentiate and change their cell types (Nakanishi et al., 2014). This has led to numerous revisions of the concept at the functional, developmental, and even molecular (gene expression) level. Here, we use the term 'cell type' as 'a classification unit based on the combined observations of a cell morphology and gene expression profile, which is driven by a gene regulatory network and can be repeatedly found within the context of a species'. These cell types can be part of either a spatially- or a temporally-integrated life cycle.
- **Cell module:** a functional unit oriented to perform a certain task in the cell. Cell modules can be formed by sets of effectors and/or regulatory proteins involved in a particular process, such as adhesion, cell-cell communication, motility, and signalling. Cell modules can be homologous in structure and (sometimes function) between different species and lineages.

BOX 2: Genomic caveats: gene loss and annotation

It is important not to forget gene loss could explain why some of these genes are not found in the most basal lineages. We know that the genomes of sponges and ctenophores underwent substantial gene loss, which could have worked as a powerful evolutionary driver leading to the highly derived organisms they are today (O'Malley et al., 2016). This makes it difficult to infer the sister group to the rest of animals, as the complexity achieved by, say, ctenophores seems counter-intuitive to the idea that they may be the sister group. There are several cases of genes not having been found in sponges and ctenophores: one example is the patchy distribution of Hox genes, the result from an intricate events of gene conservation and loss (reviewed in (Lanna, 2015)); another example are different components of the Wnt and TGF-β pathway, with missing effectors and downstream signaling components. Similar events of gene loss have led to the absence of JNK and Notch pathways in placozoans respect to sponges and ctenophores (Moroz et al., 2014; Srivastava et al., 2010, 2008). Upon comparison with the unicellular relatives of animals, it becomes evident that there were major events of gene loss at the origin of animals (Richter et al., 2018). This could hint the realization that these pathways are not working exactly the way we know them in bilaterians or vertebrates and are much more complex than we thought.

Similarly, it is important to mention that many of the genes found in these organisms are of unknown function because they don't resemble any other annotated genes, as the databases have been mostly vertebrate-centered. There is a need to study the function of these genes. All of these issues could (mis)lead us to think about the animal LCA (or the first animal) as relatively simple organisms.

BOX 3: Unicellular holozoans as emerging model organisms

Novel biological questions require suitable and phylogenetically better-placed non-traditional models, or *emerging model systems* (Goldstein and King, 2016). Unicellular holozoans are ideal systems to study the evolutionary origins of animals: phylogenetically they are the closest extant unicellular relatives of animals; they possess complex genetic repertoires of genes related to multicellular functions in animals and they present complex developmental modes with multicellular-like structures, reported to be differentially regulated at distinct levels in some of them.

S. rosetta has become a reference benchmark for gene-function studies among choanoflagellates (Goldstein and King, 2016; Hoffmeyer and Burkhardt, 2016). In fact, the first functional link between genotype and phenotype was set out using forward genetics in *S. rosetta,* notable for its experimental tractability relative to other choanoflagellate species (de Mendoza and Ruiz-Trillo, 2014; Levin et al., 2014, p. 201). A highly penetrant mutant defective in *rosette* development, named *rosetteless,* allowed the identification of a C-type lectin gene, an exclusive choanoflagellate and metazoan gene family, as essential for the formation and establishment of *rosette* colonies (Levin et al., 2014). Further genetic screens on two other *rosette* defect mutants recovered from the same study, *Jumble* and *Couscous*, which aggregate into amorphous cell clumps, mapped to genes encoding two different predicted glycosyltransferases, suggesting a pre-metazoan role of glycosyltransferases in regulating development and preventing abnormal tumor-like multicellularity (Levin et al., 2014; Wetzel et al., 2018). In recent years, the development of genetic tools in unicellular holozoans has opened new avenues of functional-based studies for investigating gene function. In *S. rosetta*, a robust transfection method for delivering and expressing transgenes using Nucleofection, an electroporation-based method for DNA delivery (Booth et al., 2018; Caro et al., 2012; Janse et al., 2006; Vinayak et al., 2015), allowed complementation of each of the mutations of *Jumble* and *Couscous* rosette defect mutants by overexpressing the wild-type alleles (Wetzel et al., 2018). Additionally, *S. rosetta* organelles can be now visualized in live cells thanks to a self-engineered panel of fluorescently-tagged endogenous proteins (Booth et al., 2018). Transfection also enabled the first *in vivo* characterization of Septins in *S. rosetta*, a class of cytoskeletal proteins which localized to the basal poles of cells, resembling Septin localization in animal epithelia (Bowen et al., 2011). Moreover, transfected cells can be enriched using a plasmid with a puromycin resistance gene (Booth et al., 2018; Wetzel et al., 2018).

Transient transfection has been developed for the filasterean *C. owczarzaki* using the classical calcium-phosphate precipitation method coupled with a glycerol shock (Parra-Acero et al., 2018). A panel of fluorescently-tagged cassettes has also allowed the visualization of several subcellular structures in *Capsaspora* live cells, which revealed with unprecedented detail the three-dimensional organisation of filopodia and actin bundles around the cell body (Parra-Acero et al., 2018).

In the ichthyosporean *C. fragrantissima,* transient transfection by electroporation has allowed tracing of nuclear divisions in a growing cell *in vivo*, showing that these divisions are strictly synchronized (Suga and Ruiz-Trillo, 2013). Moreover, two strategies for gene silencing using RNA interference by small interference RNAs (siRNA) and Morpholinos were used to successfully knock down the expression of target genes in a dosage-dependent manner (Suga and Ruiz-Trillo, 2013). This tool has been recently used to analyze the function of *c-Src* kinase animal homologs throughout the life cycle of *C. fragrantissima*

live cells, suggesting that an existing tyrosine-specific phosphatase was co-opted for the role of *Src* regulation in the highly reduced kinome of *C. fragrantissima* (Suga and Miller, 2018).

Transfection has also been developed by our colleagues for two additional unicellular holozoan species: the ichthyosporean *Abeoforma whisleri* (Casacuberta et al., unpublished) and the teretosporean *Corallochytrium limacisporum* (Kozyczkowska et al., in preparation, Casacuberta et al., unpublished). Both species can be transiently transfected with fluorescently-tagged reporter cassettes containing endogenous genes using the same strategy as in *S. rosetta.* Moreover, *C. limacisporum* transfectants can be stably maintained using antibiotic-based selection, which are further being characterized.

Finally, efforts towards establishing stable transfection using antibiotic-based selection or genome editing tools, such as the CRISPR/Cas9 technology, are underway in the choanoflagellate *S. rosetta* and the filasterean *C. owczarzaki*. Efforts are also underway to develop genetic tools in other unicellular holozoan species.

Figures

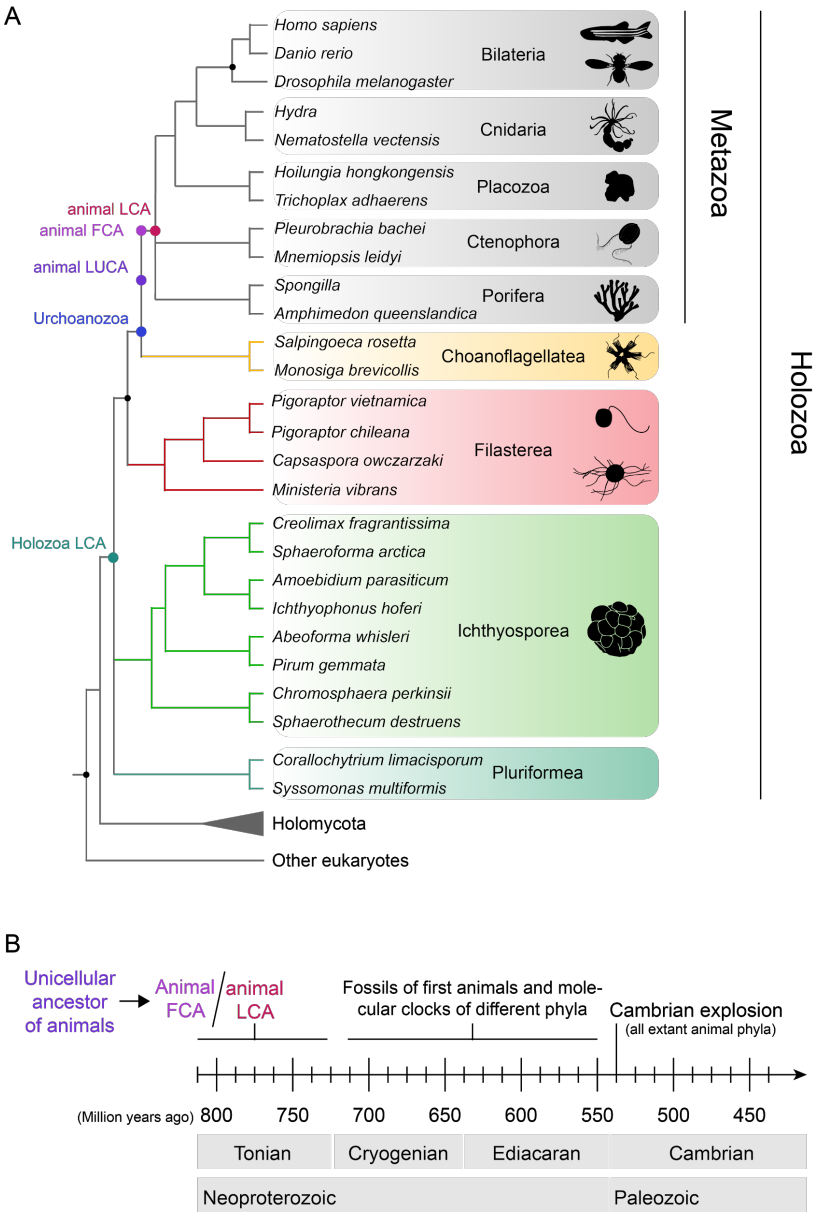


Figure 1: The origin of animals. A: Phylogenetic tree of animals and the major groups of unicellular relatives: choanoflagellates, filastereans, ichthyosporeans, and Pluriformea. Highlighted nodes indicate different ancestors that we can reconstruct and that are important to understand the transition to animal multicellularity. B: a timeline of different events during early animal evolution. The transition to animal multicellularity, and therefore the first animals, occurred some time at the end of the Tonian period, according to molecular clocks. The first fossils of recognizable animals are from the Ediacaran period, with molecular clocks extending the emergence of different animal phyla back to the cryogenian. By the time of the Cambrian, all extant animal phyla had originated.

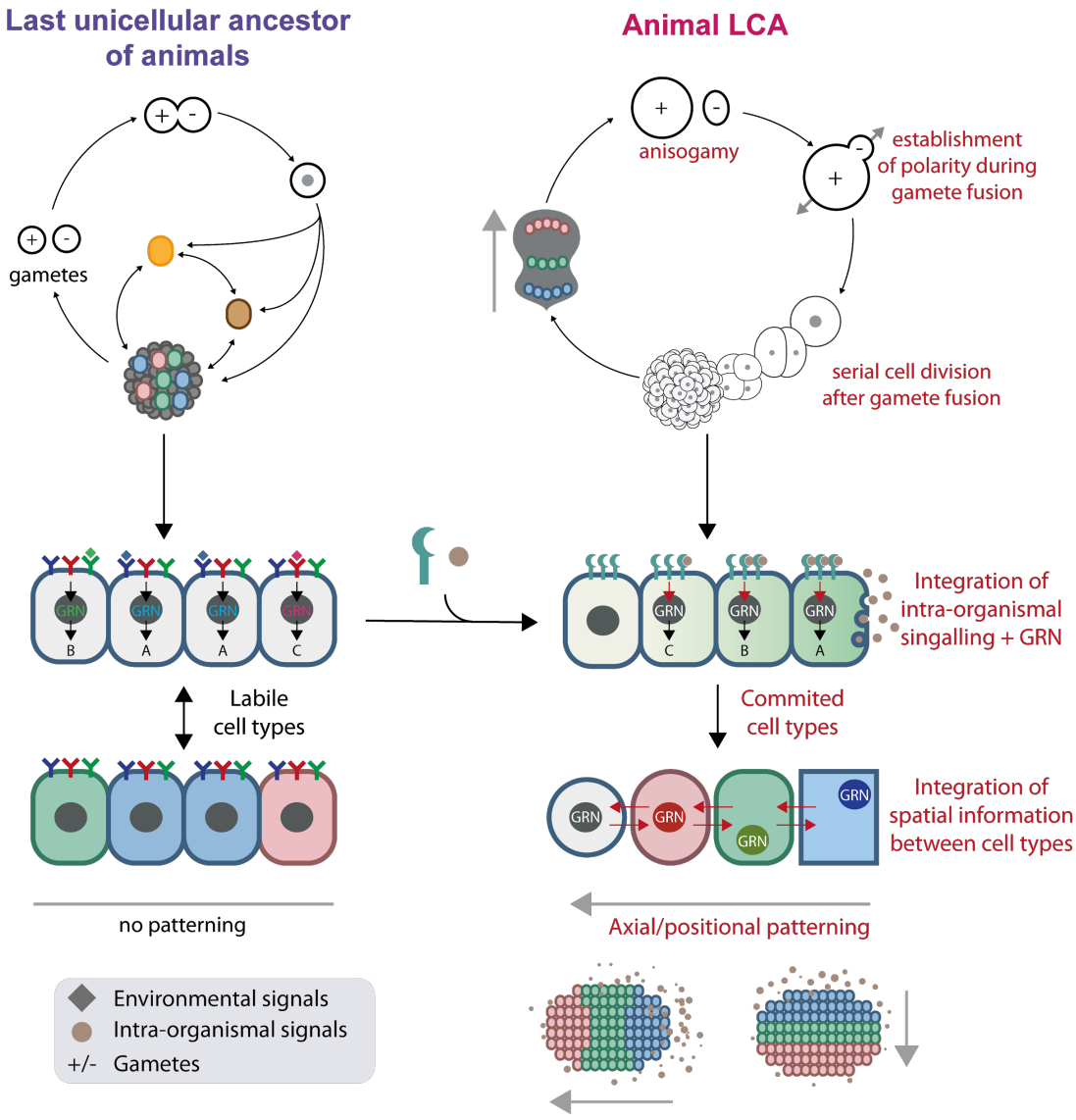


Figure 2: Important changes for the origin of animals. The single-celled ancestor of animals likely possessed a life cycle comprised of different life stages regulated temporally, plus at least one multicellular stage. Cells within this structure were able to respond to different environmental stimuli thanks to a complex repertoire of signalling molecules and GRNs. This organism was capable of sexual reproduction. The transition to animal origins likely involved some changes in this life cycle (highlighted in red), already present by the time of the animal LCA: cells within the multicellular structure acquired the ability to integrate spatial information from within the organism by making use of morphogenetic tools (such as ligands, receptors, and GRNs), which allowed the spatial organization of cell types. A greater ability to establish different cell types independently of the environment translates to the emergence of rudimentary morphogenetic plans, consisting of simple positional patterns (such as a primary axis) where different cell types localize to different regions of the organism. Concomitantly, this developmental program was conjoined to the sexual reproduction program, by which gamete

fusion was able to trigger the formation of a multicellular structure through asymmetrical, serial division. It is worth emphasizing that the visual depictions presented here are mere representations of general concepts, and that we are by no means taking positions regarding specific details, such as the real structure of the life cycles, the number of cells, genes, molecules, and regulatory networks implicated, the axial patterning, or the morphological details of these organisms.

REFERENCES

Abedin, M., King, N., 2010. Diverse evolutionary paths to cell adhesion. Trends Cell Biol. 20, 734–742. https://doi.org/10.1016/j.tcb.2010.08.002

Adamska, M., Degnan, S.M., Green, K.M., Adamski, M., Craigie, A., Larroux, C., Degnan, B.M., 2007. Wnt and TGF-β expression in the sponge Amphimedon queenslandica and the origin of metazoan embryonic patterning. PLoS ONE 2. https://doi.org/10.1371/journal.pone.0001031

Adl, S.M., Bass, D., Lane, C.E., Lukeš, J., Schoch, C.L., Smirnov, A., Agatha, S., Berney, C., Brown, M.W., Burki, F., Cárdenas, P., Čepička, I., Chistyakova, L., Campo, J. del, Dunthorn, M., Edvardsen, B., Eglit, Y., Guillou, L., Hampl, V., Heiss, A.A., Hoppenrath, M., James, T.Y., Karnkowska, A., Karpov, S., Kim, E., Kolisko, M., Kudryavtsev, A., Lahr, D.J.G., Lara, E., Gall, L.L., Lynn, D.H., Mann, D.G., Massana, R., Mitchell, E.A.D., Morrow, C., Park, J.S., Pawlowski, J.W., Powell, M.J., Richter, D.J., Rueckert, S., Shadwick, L., Shimano, S., Spiegel, F.W., Torruella, G., Youssef, N., Zlatogursky, V., Zhang, Q., 2019. Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. J. Eukaryot. Microbiol. 66, 4–119. https://doi.org/10.1111/jeu.12691

Ajduk, A., Zernicka-Goetz, M., 2016. Polarity and cell division orientation in the cleavage embryo: From worm to human. Mol. Hum. Reprod. 22, 691–703. https://doi.org/10.1093/molehr/gav068

Alegado, R.A., Brown, L.W., Cao, S., Dermenjian, R.K., Zuzow, R., Fairclough, S.R., Clardy, J., King, N., 2012. A bacterial sulfonolipid triggers multicellular development in the closest living relatives of animals. eLife 1, e00013. https://doi.org/10.7554/eLife.00013

Alegado, R.A., King, N., 2014. Bacterial Influences on Animal Origins. Cold Spring Harb. Perspect. Biol. 6, a016162–a016162. https://doi.org/10.1101/cshperspect.a016162

Antcliffe, J.B., Callow, R.H.T., Brasier, M.D., 2014. Giving the early fossil record of sponges a squeeze. Biol. Rev. 89, 972–1004. https://doi.org/10.1111/brv.12090

Aravind, L., Subramanian, G., 1999. Origin of multicellular eukaryotes – insights from proteome comparisons. Curr. Opin. Genet. Dev. 9, 688–694. https://doi.org/10.1016/S0959-437X(99)00028-3

Arendt, D., Bertucci, P.Y., Achim, K., Musser, J.M., 2019. Evolution of neuronal types

and families. Curr. Opin. Neurobiol. 56, 144–152. https://doi.org/10.1016/j.conb.2019.01.022

Arndt, H., Dietrich, D., Auer, B., Cleven, E.-J., Gräfenhan, T., Weitere, M., Mylnikov, A.P., 2000. Functional diversity of heterotrophic flagellates in aquatic ecosystems 29.

Arroyo, A.S., López-Escardó, D., Kim, E., Ruiz-Trillo, I., Najle, S.R., 2018. Novel Diversity of Deeply Branching Holomycota and Unicellular Holozoans Revealed by Metabarcoding in Middle Paraná River, Argentina. Front. Ecol. Evol. 6, 99. https://doi.org/10.3389/fevo.2018.00099

Baran, Y., Sebe-pedros, A., Lubling, Y., Giladi, A., Chomsky, E., Meir, Z., 2018. MetaCell: analysis of single cell RNA-seq data using k-NN graph partitions. BioRxiv Bioinforma. https://doi.org/10.1101/437665

Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., Kim, T., Misquitta-Ali, C.M., Wilson, M.D., Kim, P.M., Odom, D.T., Frey, B.J., Blencowe, B.J., 2012. The Evolutionary Landscape of Alternative Splicing in Vertebrate Species. Science 338, 1587–1593. https://doi.org/10.1126/science.1230612

Belahbib, H., Renard, E., Santini, S., Jourda, C., Claverie, J.-M., Borchiellini, C., Le Bivic, A., 2018. New genomic data and analyses challenge the traditional vision of animal epithelium evolution. BMC Genomics 19, 393. https://doi.org/10.1186/s12864-018-4715-9

Benny, G.L., O'Donnell, K., 2000. Amoebidium parasiticum Is a Protozoan, Not a Trichomycete. Mycologia 92, 1133. https://doi.org/10.2307/3761480

Bobrovskiy, I., Hope, J.M., Ivantsov, A., Nettersheim, B.J., Hallmann, C., Brocks, J.J., 2018. Ancient steroids establish the Ediacaran fossil Dickinsonia as one of the earliest animals. Science 1249, 1246–1249.

Bonner, J.T., 1998. The origins of multicellularity. Integr. Biol. 27–36. https://doi.org/doi.org/10.1002/(SICI)1520-6602(1998)1:1<27::AID-INBI4>3.0.CO;2-6

Booth, D.S., Szmidt-Middleton, H., King, N., 2018. Transfection of choanoflagellates illuminates their cell biology and the ancestry of animal septins. Mol. Biol. Cell 29, 3026–3038. https://doi.org/10.1091/mbc.E18-08-0514

Boraas, M.E., Seale, D.B., Boxhorn, J.E., 1998. Phagotrophy by flagellate selects for colonial prey: A possible origin of multicellularity. Evol. Ecol. 12, 153–164. https://doi.org/10.1023/A:1006527528063

Bowen, J.R., Hwang, D., Bai, X., Roy, D., Spiliotis, E.T., 2011. Septin GTPases spatially guide microtubule organization and plus end dynamics in polarizing epithelia. J. Cell Biol. 194, 187–197. https://doi.org/10.1083/jcb.201102076

Bråte, J., Neumann, R.S., Fromm, B., Haraldsen, A.A.B., Tarver, J.E., Suga, H., Donoghue, P.C.J., Peterson, K.J., Ruiz-Trillo, I., Grini, P.E., Shalchian-Tabrizi, K., 2018. Unicellular Origin of the Animal MicroRNA Machinery. Curr. Biol. 28, 3288-3295.e5. https://doi.org/10.1016/j.cub.2018.08.018

Brennan, J.J., Messerschmidt, J.L., Williams, L.M., Matthews, B.J., Reynoso, M., Gilmore, T.D., 2017. Sea anemone model has a single Toll-like receptor that can function in pathogen detection, NF-κB signal transduction, and development. Proc. Natl. Acad. Sci. 114, E10122–E10131. https://doi.org/10.1073/pnas.1711530114

Brown, M.W., Kolisko, M., Silberman, J.D., Roger, A.J., 2012. Aggregative Multicellularity Evolved Independently in the Eukaryotic Supergroup Rhizaria. Curr. Biol. 22, 1123–1127. https://doi.org/10.1016/j.cub.2012.04.021

Brown, M.W., Silberman, J.D., 2013. The Non-dictyostelid Sorocarpic Amoebae, in: Romeralo, M., Baldauf, S., Escalante, R. (Eds.), Dictyostelids: Evolution, Genomics and Cell Biology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 219–242. https://doi.org/10.1007/978-3-642-38487-5_12

Brown, M.W., Spiegel, F.W., Silberman, J.D., 2009. Phylogeny of the “Forgotten” Cellular Slime Mold, Fonticula alba, Reveals a Key Evolutionary Branch within Opisthokonta. Mol. Biol. Evol. 26, 2699–2709. https://doi.org/10.1093/molbev/msp185

Brunet, T., King, N., 2017. The Origin of Animal Multicellularity and Cell Differentiation. Dev. Cell 43, 124–140. https://doi.org/10.1016/j.devcel.2017.09.016

Brunet, T., Larson, B.T., Linden, T.A., Vermeij, M.J.A., McDonald, K., King, N., 2019. Light-regulated collective contractility in a multicellular choanoflagellate (preprint). Cell Biology. https://doi.org/10.1101/661009

Bucher, D., Anderson, P.A.V., 2015. Evolution of the first nervous systems - What can we surmise? J. Exp. Biol. 218, 501–503. https://doi.org/10.1242/jeb.111799

Budd, G.E., Jensen, S., 2017. The origin of the animals and a ‘Savannah’’ hypothesis for early bilaterian evolution.’ Biol. Rev. 92, 446–473. https://doi.org/10.1111/brv.12239

Burton, P.M., 2008. Insights from diploblasts; the evolution of mesoderm and muscle. https://doi.org/10.1002/jez.b.21150

Butterfield, N.J., 2011. Terminal Developments in Ediacaran Embryology. Science 334, 1655–1656. https://doi.org/10.1126/science.1216125

Cafaro, M.J., 2005. Eccrinales (Trichomycetes) are not fungi, but a clade of protists at the early divergence of animals and fungi. Mol. Phylogenet. Evol. 35, 21–34. https://doi.org/10.1016/j.ympev.2004.12.019

Cámara, M., Hardman, A., Williams, P., Milton, D., 2002. Quorum sensing in Vibrio cholerae. Nat. Genet. 32, 217.

Canfield, D.E., Poulton, S.W., Knoll, A.H., Narbonne, G.M., Ross, G., Goldberg, T., Strauss, H., 2008. Ferruginous conditions dominated later neoproterozoic deep-water chemistry. Science 321, 949–952. https://doi.org/10.1126/science.1154499

Canfield, D.E., Poulton, S.W., Narbonne, G.M., 2007. Late-Neoproterozoic deep-ocean oxygenation and the rise of animal life. Science 315, 92–95. https://doi.org/10.1126/science.1135013

Cao, L., Chen, F., Yang, X., Xu, W., Xie, J., Yu, L., 2014. Phylogenetic analysis of CDK and cyclin proteins in premetazoan lineages. BMC Evol. Biol. 14, 10. https://doi.org/10.1186/1471-2148-14-10

Caro, F., Miller, M.G., DeRisi, J.L., 2012. Plate-based transfection and culturing technique for genetic manipulation of Plasmodium falciparum. Malar. J. 11, 22. https://doi.org/10.1186/1475-2875-11-22

Carr, M., Leadbeater, B.S.C., Hassan, R., Nelson, M., Baldauf, S.L., 2008. Molecular phylogeny of choanoflagellates, the sister group to Metazoa. Proc. Natl. Acad. Sci. 105, 16641–16646. https://doi.org/10.1073/pnas.0801667105

Carr, M., Richter, D.J., Fozouni, P., Smith, T.J., Jeuck, A., Leadbeater, B.S.C., Nitsche, F., 2017. A six-gene phylogeny provides new insights into choanoflagellate evolution. Mol. Phylogenet. Evol. 107, 166–178. https://doi.org/10.1016/j.ympev.2016.10.011

Cavalier-Smith, T., 2017. Origin of animal multicellularity: precursors, causes, consequences—the choanoflagellate/sponge transition, neurogenesis and the Cambrian explosion. Philos. Trans. R. Soc. B Biol. Sci. 372, 20150476. https://doi.org/10.1098/rstb.2015.0476

Cavalier-Smith, T., 1987. The Origin of Eukaryote and Archaebacterial Cells. Ann. N. Y. Acad. Sci. 503, 17–54. https://doi.org/10.1111/j.1749-6632.1987.tb40596.x

Cavalier-Smith, T., Chao, E.E.Y., 2003. Phylogeny and classification of phylum Cercozoa (Protozoa). Protist 154, 341–358.

https://doi.org/10.1078/143446103322454112

Chang, C., Hemmati-Brivanlou, A., 1998. Cell fate determination in embryonic ectoderm. https://doi.org/10.1002/(SICI)1097-4695(199808)36:2<128::AID-NEU3>3.0.CO;2-3

Chang, E.S., Neuhof, M., Rubinstein, N.D., Diamant, A., Philippe, H., Huchon, D., Cartwright, P., 2015. Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. Proc. Natl. Acad. Sci. U. S. A. 112, 14912–14917. https://doi.org/10.1073/pnas.1511468112

Chapman, J.A., Kirkness, E.F., Simakov, O., Hampson, S.E., Mitros, T., Weinmaier, T., Rattei, T., Balasubramanian, P.G., Borman, J., Busam, D., Disbennett, K., Pfannkoch, C., Sumin, N., Sutton, G.G., Viswanathan, L.D., Walenz, B., Goodstein, D.M., Hellsten, U., Kawashima, T., Prochnik, S.E., Putnam, N.H., Shu, S., Blumberg, B., Dana, C.E., Gee, L., Kibler, D.F., Law, L., Lindgens, D., Martinez, D.E., Peng, J., Wigge, P.A., Bertulat, B., Guder, C., Nakamura, Y., Ozbek, S., Watanabe, H., Khalturin, K., Hemmrich, G., Franke, A., Augustin, R., Fraune, S., Hayakawa, E., Hayakawa, S., Hirose, M., Hwang, J.S., Ikeo, K., Nishimiya-Fujisawa, C., Ogura, A., Takahashi, T., Steinmetz, P.R.H., Zhang, X., Aufschnaiter, R., Eder, M.K., Gorny, A.K., Salvenmoser, W., Heimberg, A.M., Wheeler, B.M., Peterson, K.J., Böttger, A., Tischler, P., Wolf, A., Gojobori, T., Remington, K.A., Strausberg, R.L., Venter, J.C., Technau, U., Hobmayer, B., Bosch, T.C.G., Holstein, T.W., Fujisawa, T., Bode, H.R., David, C.N., Rokhsar, D.S., Steele, R.E., 2010. The dynamic genome of Hydra. Nature 464, 592–596. https://doi.org/10.1038/nature08830

Chen, L., Xiao, S., Pang, K., Zhou, C., Yuan, X., 2014. Cell differentiation and germ–soma separation in Ediacaran animal embryo-like fossils. Nature 516, 238–241. https://doi.org/10.1038/nature13766

Clark, H.J., 1866. Conclusive proofs of the animality of the ciliate sponges, and of their affinities with the Infusoria flagellata. Am. J. Sci. s2-42, 320–324. https://doi.org/10.2475/ajs.s2-42.126.320

Dayel, M.J., Alegado, R.A., Fairclough, S.R., Levin, T.C., Nichols, S.A., McDonald, K., King, N., 2011. Cell differentiation and morphogenesis in the colony-forming choanoflagellate Salpingoeca rosetta. Dev. Biol. 357, 73–82. https://doi.org/10.1016/j.ydbio.2011.06.003

de Mendoza, A., Ruiz-Trillo, I., 2014. Forward genetics for back-in-time questions. eLife 3, e05218. https://doi.org/10.7554/eLife.05218

de Mendoza, A., Sebé-Pedrós, A., 2019. Origin and evolution of eukaryotic transcription factors. Curr. Opin. Genet. Dev. 58–59, 25–32. https://doi.org/10.1016/j.gde.2019.07.010

de Mendoza, A., Sebe-Pedros, A., Sestak, M.S., Matejcic, M., Torruella, G., Domazet-Loso, T., Ruiz-Trillo, I., 2013. Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. Proc. Natl. Acad. Sci. 110, E4858–E4866. https://doi.org/10.1073/pnas.1311818110

de Mendoza, A., Suga, H., Permanyer, J., Irimia, M., Ruiz-Trillo, I., 2015. Complex transcriptional regulation and independent evolution of fungal-like traits in a relative of animals. eLife 4, e08904. https://doi.org/10.7554/eLife.08904

del Campo, J., Guillou, L., Hehenberger, E., Logares, R., López-García, P., Massana, R., 2016. Ecological and evolutionary significance of novel protist lineages. Eur. J. Protistol. 55, 4–11. https://doi.org/10.1016/j.ejop.2016.02.002

del Campo, J., Mallo, D., Massana, R., de Vargas, C., Richards, T.A., Ruiz-Trillo, I., 2015. Diversity and distribution of unicellular opisthokonts along the European coast analysed using high-throughput sequencing: Diversity of unicellular opisthokonts through HTS. Environ. Microbiol. 17, 3195–3207. https://doi.org/10.1111/1462-2920.12759

del Campo, J., Massana, R., 2011. Emerging Diversity within Chrysophytes, Choanoflagellates and Bicosoecids Based on Molecular Surveys. Protist 162, 435–448. https://doi.org/10.1016/j.protis.2010.10.003

del Campo, J., Ruiz-Trillo, I., 2013. Environmental Survey Meta-analysis Reveals Hidden Diversity among Unicellular Opisthokonts. Mol. Biol. Evol. 30, 802–805. https://doi.org/10.1093/molbev/mst006

Denbo, S., Aono, K., Kai, T., Yagasaki, R., Ruiz-Trillo, I., Suga, H., 2019. Revision of the *Capsaspora* genome using read mating information adjusts the view on premetazoan genome. Dev. Growth Differ. 61, 34–42. https://doi.org/10.1111/dgd.12587

Dohrmann, M., Wörheide, G., 2017. Dating early animal evolution using phylogenomic data/631/181/735/631/181/414 article. Sci. Rep. 7, 1–6. https://doi.org/10.1038/s41598-017-03791-w

dos Reis, M., Thawornwattana, Y., Angelis, K., Telford, M.J., Donoghue, P.C.J., Yang, Z., 2015. Uncertainty in the Timing of Origin of Animals and the Limits of Precision in Molecular Timescales. Curr. Biol. 25, 2939–2950. https://doi.org/10.1016/j.cub.2015.09.066

Douzery, E.J.P., Snell, E.A., Bapteste, E., Delsuc, F., Philippe, H., 2004. The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? Proc. Natl. Acad. Sci. U. S. A. 101, 15386–15391. https://doi.org/10.1073/pnas.0403984101

Du, Q., Kawabe, Y., Schilde, C., Chen, Z., Schaap, P., 2015. The Evolution of Aggregative Multicellularity and Cell–Cell Communication in the Dictyostelia. J. Mol. Biol. 427, 3722–3733. https://doi.org/10.1016/j.jmb.2015.08.008

Dudin, O., Ondracka, A., Grau-Bové, X., A.B. Haraldsen, A., Toyoda, A., Suga, H., Bråte, J., Ruiz-Trillo, I., 2019. A unicellular relative of animals generates an epithelium-like cell layer by actomyosin-dependent cellularization (preprint). Evolutionary Biology. https://doi.org/10.1101/563726

Elliott, T.A., Gregory, T.R., 2015. What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. Philos. Trans. R. Soc. B Biol. Sci. 370, 20140331. https://doi.org/10.1098/rstb.2014.0331

Fahey, B., Degnan, B.M., 2010. Origin of animal epithelia: insights from the sponge genome: Evolution of epithelia. Evol. Dev. 12, 601–617. https://doi.org/10.1111/j.1525-142X.2010.00445.x

Fairclough, S.R., Chen, Z., Kramer, E., Zeng, Q., Young, S., Robertson, H.M., Begovic, E., Richter, D.J., Russ, C., Westbrook, M.J., Manning, G., Lang, B.F., Haas, B., Nusbaum, C., King, N., 2013. Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate Salpingoeca rosetta. Genome Biol. 14, R15. https://doi.org/10.1186/gb-2013-14-2-r15

Fairclough, S.R., Dayel, M.J., King, N., 2010. Multicellular Development in a Choanoflagellate. Curr. Biol. CB 20, R875–R876. https://doi.org/10.1016/j.cub.2010.09.014

Fatema, U., Ali, M.F., Hu, Z., Clark, A.J., Kawashima, T., 2019. Gamete nuclear migration in animals and plants. Front. Plant Sci. 10, 1–14. https://doi.org/10.3389/fpls.2019.00517

Feldman, S.H., Wimsatt, J.H., Green, D.E., 2005. Phylogenetic classification of the frog pathogen Amphibiothecum (Dermosporidium) penneri based on small ribosomal subunit sequencing. J. Wildl. Dis. 41, 701–706. https://doi.org/10.7589/0090-3558-41.4.701

Fidler, A.L., Darris, C.E., Chetyrkin, S.V., Pedchenko, V.K., Boudko, S.P., Brown, K.L., Gray Jerome, W., Hudson, J.K., Rokas, A., Hudson, B.G., 2017. Collagen IV and basement membrane at the evolutionary dawn of metazoan tissues. eLife

6, e24176. https://doi.org/10.7554/eLife.24176

Gauthier, M.E.A., Du Pasquier, L., Degnan, B.M., 2010. The genome of the sponge Amphimedon queenslandica provides new perspectives into the origin of Toll-like and interleukin 1 receptor pathways. Evol. Dev. 12, 519–533. https://doi.org/10.1111/j.1525-142X.2010.00436.x

Glockling, S.L., Marshall, W.L., Gleason, F.H., 2013. Phylogenetic interpretations and ecological potentials of the Mesomycetozoea (Ichthyosporea). Fungal Ecol. 6, 237–247. https://doi.org/10.1016/j.funeco.2013.03.005

Gold, D.A., Katsuki, T., Li, Y., Yan, X., Regulski, M., Ibberson, D., Holstein, T., Steele, R.E., Jacobs, D.K., Greenspan, R.J., 2019. The genome of the jellyfish Aurelia and the evolution of animal complexity. Nat. Ecol. Evol. 3, 96–104. https://doi.org/10.1038/s41559-018-0719-8

Goldstein, B., King, N., 2016. The Future of Cell Biology: Emerging Model Organisms. Trends Cell Biol. 26, 818–824. https://doi.org/10.1016/j.tcb.2016.08.005

Grau-Bové, X., Torruella, G., Donachie, S., Suga, H., Leonard, G., Richards, T.A., Ruiz-Trillo, I., 2017. Dynamics of genomic innovation in the unicellular ancestry of animals. eLife 6, e26036. https://doi.org/10.7554/eLife.26036

Grosberg, R.K., Strathmann, R.R., 2007. The Evolution of Multicellularity: A Minor Major Transition? Annu. Rev. Ecol. Evol. Syst. 38, 621–654. https://doi.org/10.1146/annurev.ecolsys.36.102403.114735

Hachet, O., Berthelot-Grosjean, M., Kokkoris, K., Vincenzetti, V., Moosbrugger, J., Martin, S.G., 2011. A Phosphorylation Cycle Shapes Gradients of the DYRK Family Kinase Pom1 at the Plasma Membrane. Cell 145, 1116–1128. https://doi.org/10.1016/j.cell.2011.05.014

Haeckel, E., 1874. Anthropogenie : oder, Entwickelungsgeschichte des Menschen. Engelmann, Leipzig : https://doi.org/10.5962/bhl.title.119035

Hassett, B.T., López, J.A., Gradinger, R., 2015. Two New Species of Marine Saprotrophic Sphaeroformids in the Mesomycetozoea Isolated From the Sub-Arctic Bering Sea. Protist 166, 310–322. https://doi.org/10.1016/j.protis.2015.04.004

Hazkani-Covo, E., Levanon, E.Y., Rotman, G., Graur, D., Novik, A., 2004. Evolution of multicellularity in Metazoa: Comparative analysis of the subcellular localization of proteins in Saccharomyces, Drosophila and Caenorhabditis. Cell Biol. Int. 28, 171–178. https://doi.org/10.1016/j.cellbi.2003.11.016

Hedges, S.B., Blair, J.E., Venturi, M.L., Shoe, J.L., 2004. A molecular timescale of

eukaryote evolution and the rise of complex multicellular life. BMC Evol. Biol. 4. https://doi.org/10.1186/1471-2148-4-2

Heger, T.J., Giesbrecht, I.J.W., Gustavsen, J., del Campo, J., Kellogg, C.T.E., Hoffman, K.M., Lertzman, K., Mohn, W.W., Keeling, P.J., 2018. High-throughput environmental sequencing reveals high diversity of litter and moss associated protist communities along a gradient of drainage and tree productivity: Moss and litter-associated protists. Environ. Microbiol. 20, 1185–1203. https://doi.org/10.1111/1462-2920.14061

Hehenberger, E., Tikhonenkov, D.V., Kolisko, M., del Campo, J., Esaulov, A.S., Mylnikov, A.P., Keeling, P.J., 2017. Novel Predators Reshape Holozoan Phylogeny and Reveal the Presence of a Two-Component Signaling System in the Ancestor of Animals. Curr. Biol. 27, 2043-2050.e6. https://doi.org/10.1016/j.cub.2017.06.006

Herron, M.D., Borin, J.M., Boswell, J.C., Walker, J., Chen, I.C.K., Knox, C.A., Boyd, M., Rosenzweig, F., Ratcliff, W.C., 2019. De novo origins of multicellularity in response to predation. Sci. Rep. 9, 2328. https://doi.org/10.1038/s41598-019-39558-8

Hobmayer, B., Rentzsch, F., Kuhn, K., Happel, C.M., Von Laue, C.C., Snyder, P., Rothbächer, U., Holstein, T.W., 2000. WNT signalling molecules act in axis formation in the diploblastic metazoan Hydra. Nature 407, 186–189. https://doi.org/10.1038/35025063

Hoffmeyer, T.T., Burkhardt, P., 2016. Choanoflagellate models — Monosiga brevicollis and Salpingoeca rosetta. Curr. Opin. Genet. Dev. 39, 42–47. https://doi.org/10.1016/j.gde.2016.05.016

Irimia, M., Rukov, J.L., Roy, S.W., Vinther, J., Garcia-Fernandez, J., 2009. Quantitative regulation of alternative splicing in evolution and development. BioEssays 31, 40–50. https://doi.org/10.1002/bies.080092

Janse, C.J., Ramesar, J., Waters, A.P., 2006. High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite Plasmodium berghei. Nat. Protoc. 1, 346–356. https://doi.org/10.1038/nprot.2006.53

Josephson, R.K., 2004. Electrical properties of an excitable epithelium. J. Gen. Physiol. 74, 213–236. https://doi.org/10.1085/jgp.74.2.213

King, N., 2005. Choanoflagellates. Curr. Biol. CB 15, R113-114. https://doi.org/10.1016/j.cub.2005.02.004

King, N., 2004. The unicellular ancestry of animal development. Dev. Cell 7, 313–325.

King, N., Rokas, A., 2017. Embracing Uncertainty in Reconstructing Early Animal Evolution. Curr. Biol. 27, R1081–R1088. https://doi.org/10.1016/j.cub.2017.08.054

King, N., Westbrook, M.J., Young, S.L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I., Marr, M., Pincus, D., Putnam, N., Rokas, A., Wright, K.J., Zuzow, R., Dirks, W., Good, M., Goodstein, D., Lemons, D., Li, W., Lyons, J.B., Morris, A., Nichols, S., Richter, D.J., Salamov, A., Sequencing, J., Bork, P., Lim, W.A., Manning, G., Miller, W.T., McGinnis, W., Shapiro, H., Tjian, R., Grigoriev, I.V., Rokhsar, D., 2008. The genome of the choanoflagellate Monosiga brevicollis and the origin of metazoans. Nature 451, 783–788. https://doi.org/10.1038/nature06617

Kiss, Á.K., Ács, É., Kiss, K.T., Török, J.K., 2009. Structure and seasonal dynamics of the protozoan community (heterotrophic flagellates, ciliates, amoeboid protozoa) in the plankton of a large river (River Danube, Hungary). Eur. J. Protistol. 45, 121–138. https://doi.org/10.1016/j.ejop.2008.08.002

Klimovich, A.V., Bosch, T.C.G., 2018. Rethinking the Role of the Nervous System: Lessons From the Hydra Holobiont. BioEssays 40, 1–9. https://doi.org/10.1002/bies.201800060

Knoll, A.H., 2011. The Multiple Origins of Complex Multicellularity. Annu. Rev. Earth Planet. Sci. 39, 217–239. https://doi.org/10.1146/annurev.earth.031208.100209

Knoll, A.H., Carroll, S.B., 1999. Early animal evolution: emerging views from comparative biology and geology. Science 284, 2129–2137. https://doi.org/10.1126/science.284.5423.2129

Knoll, A.H., Hewitt, D., 2013. Phylogenetic, Functional, and Geological Perspectives on Complex Multicellularity, in: The Major Transitions in Evolution Revisited. pp. 251–270. https://doi.org/10.7551/mitpress/9780262015240.003.0013

Lang, B.F., O'Kelly, C., Nerad, T., Gray, M.W., Burger, G., 2002. The Closest Unicellular Relatives of Animals. Curr. Biol. 12, 1773–1778. https://doi.org/10.1016/S0960-9822(02)01187-9

Lanna, E., 2015. Evo-devo of non-bilaterian animals. https://doi.org/10.1590/S1415-475738320150005

Larroux, C., Fahey, B., Degnan, S.M., Adamski, M., Rokhsar, D.S., Degnan, B.M., 2007. The NK Homeobox Gene Cluster Predates the Origin of Hox Genes. Curr. Biol. 17, 706–710. https://doi.org/10.1016/j.cub.2007.03.008

Larroux, C., Fahey, B., Liubicich, D., Hinman, V.F., Gauthier, M., Gongora, M., Green, K., Wörheide, G., Leys, S.P., Degnan, B.M., 2006. Developmental expression of transcription factor genes in a demosponge: Insights into the origin of metazoan multicellularity. Evol. Dev. 8, 150–173. https://doi.org/10.1111/j.1525-142X.2006.00086.x

Larroux, C., Luke, G.N., Koopman, P., Rokhsar, D.S., Shimeld, S.M., Degnan, B.M., 2008. Genesis and Expansion of Metazoan Transcription Factor Gene Classes. Mol. Biol. Evol. 25, 980–996. https://doi.org/10.1093/molbev/msn047

Lasek-Nesselquist, E., Katz, L.A., 2001. Phylogenetic Position of Sorogena stoianovitchae and Relationships within the Class Colpodea (Ciliophora) Based on SSU rDNA Sequences. J. Eukaryot. Microbiol. 48, 604–607. https://doi.org/10.1111/j.1550-7408.2001.tb00197.x

Laumer, C.E., Fernández, R., Lemer, S., Combosch, D., Kocot, K.M., Riesgo, A., Andrade, S.C.S., Sterrer, W., Sørensen, M.V., Giribet, G., 2019. Revisiting metazoan phylogeny with genomic sampling of all phyla. Proc. R. Soc. B Biol. Sci. 286, 20190831. https://doi.org/10.1098/rspb.2019.0831

Laundon, D., Larson, B.T., McDonald, K., King, N., Burkhardt, P., 2019. The architecture of cell differentiation in choanoflagellates and sponge choanocytes 22.

Leadbeater, B.S.C., 2015. The Choanoflagellates: Evolution, Biology and Ecology by Barry S. C. Leadbeater. Cambridge University Press.

Leclère, L., Horin, C., Chevalier, S., Lapébie, P., Dru, P., Peron, S., Jager, M., Condamine, T., Pottin, K., Romano, S., Steger, J., Sinigaglia, C., Barreau, C., Quiroga Artigas, G., Ruggiero, A., Fourrage, C., Kraus, J.E.M., Poulain, J., Aury, J.M., Wincker, P., Quéinnec, E., Technau, U., Manuel, M., Momose, T., Houliston, E., Copley, R.R., 2019. The genome of the jellyfish Clytia hemisphaerica and the evolution of the cnidarian life-cycle. Nat. Ecol. Evol. 3, 801–810. https://doi.org/10.1038/s41559-019-0833-2

Lee, P.N., Pang, K., Matus, D.Q., Martindale, M.Q., 2006. A WNT of things to come: Evolution of Wnt signaling and polarity in cnidarians. https://doi.org/10.1016/j.semcdb.2006.05.002

Leininger, S., Adamski, M., Bergum, B., Guder, C., Liu, J., Laplante, M., Br\a ate, J., Hoffmann, F., Fortunato, S., Jordal, S., Rapp, H.T., Adamska, M., 2014. Developmental gene expression provides clues to relationships between sponge and eumetazoan body plans. Nat. Commun. 5.

https://doi.org/10.1038/ncomms4905

Levin, T.C., Greaney, A.J., Wetzel, L., King, N., 2014. The rosetteless gene controls development in the choanoflagellate S. rosetta. eLife 3, e04070. https://doi.org/10.7554/eLife.04070

Levin, T.C., King, N., 2013. Evidence for Sex and Recombination in the Choanoflagellate Salpingoeca rosetta. Curr. Biol. 23, 2176–2180. https://doi.org/10.1016/j.cub.2013.08.061

Lewis, S.L., Tam, P.P.L., 2006. Definitive endoderm of the mouse embryo: Formation, cell fates, and morphogenetic function. Dev. Dyn. 235, 2315–2329. https://doi.org/10.1002/dvdy.20846

Liu, M., Walch, H., Wu, S., Grigoriev, A., 2005. Significant expansion of exon-bordering protein domains during animal proteome evolution. Nucleic Acids Res. 33, 95–105. https://doi.org/10.1093/nar/gki152

Lohr, J.N., Laforsch, C., Koerner, H., Wolinska, J., 2010. A Daphnia parasite (Caullerya mesnili) constitutes a new member of the Ichthyosporea, a group of protists near the animal-fungi divergence. J. Eukaryot. Microbiol. 57, 328–336. https://doi.org/10.1111/j.1550-7408.2010.00479.x

López-Escardó, D., Paps, J., de Vargas, C., Massana, R., Ruiz-Trillo, I., del Campo, J., 2018. Metabarcoding analysis on European coastal samples reveals new molecular metazoan diversity. Sci. Rep. 8, 9106. https://doi.org/10.1038/s41598-018-27509-8

Lord, J.C., Hartzer, K.L., Kambhampati, S., 2012. A nuptially transmitted ichthyosporean symbiont of Tenebrio molitor (Coleoptera: Tenebrionidae). J. Eukaryot. Microbiol. 59, 246–250. https://doi.org/10.1111/j.1550-7408.2012.00617.x

Love, G.D., Grosjean, E., Stalvies, C., Fike, D.A., Grotzinger, J.P., Bradley, A.S., Kelly, A.E., Bhatia, M., Meredith, W., Snape, C.E., Bowring, S.A., Condon, D.J., Summons, R.E., 2009. Fossil steroids record the appearance of Demospongiae during the Cryogenian period. Nature 457, 718–721. https://doi.org/10.1038/nature07673

Lyons, T.W., Reinhard, C.T., Planavsky, N.J., 2014. The rise of oxygen in Earth's early ocean and atmosphere. Nature 506, 307–315. https://doi.org/10.1038/nature13068

Mackie, G.O., 2004. Epithelial Conduction in Hydromedusae. J. Gen. Physiol. 52, 600–621. https://doi.org/10.1085/jgp.52.3.600

Manning, G., Young, S.L., Miller, W.T., Zhai, Y., 2008. The protist, Monosiga brevicollis, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan. Proc. Natl. Acad. Sci. 105, 9674–9679. https://doi.org/10.1073/pnas.0801314105

Marioni, J.C., Arendt, D., 2017. How Single-Cell Genomics Is Changing Evolutionary and Developmental Biology. Annu. Rev. Cell Dev. Biol. 33, 537–553. https://doi.org/10.1146/annurev-cellbio-100616-060818

Marshall, W.L., Berbee, M.L., 2011. Facing Unknowns: Living Cultures (Pirum gemmata gen. nov., sp. nov., and Abeoforma whisleri, gen. nov., sp. nov.) from Invertebrate Digestive Tracts Represent an Undescribed Clade within the Unicellular Opisthokont Lineage Ichthyosporea (Mesomycetozoea). Protist 162, 33–57. https://doi.org/10.1016/j.protis.2010.06.002

Marshall, W.L., Celio, G., McLaughlin, D.J., Berbee, M.L., 2008. Multiple Isolations of a Culturable, Motile Ichthyosporean (Mesomycetozoa, Opisthokonta), Creolimax fragrantissima n. gen., n. sp., from Marine Invertebrate Digestive Tracts. Protist 159, 415–433. https://doi.org/10.1016/j.protis.2008.03.003

Martin, S.G., Berthelot-Grosjean, M., 2009. Polar gradients of the DYRK-family kinase Pom1 couple cell length with the cell cycle. Nature 459, 852–856. https://doi.org/10.1038/nature08054

Martindale, M.Q., Henry, J.Q., 1999. Intracellular fate mapping in a basal metazoan, the ctenophore Mnemiopsis leidyi, reveals the origins of mesoderm and the existence of indeterminate cell lineages. Dev. Biol. 214, 243–257. https://doi.org/10.1006/dbio.1999.9427

Medina, E.M., Turner, J.J., Gordân, R., Skotheim, J.M., Buchler, N.E., 2016. Punctuated evolution and transitional hybrid network in an ancestral cell cycle of fungi. eLife 5. https://doi.org/10.7554/eLife.09492

Mendoza, L., Taylor, J.W., Ajello, L., 2002. The Class Mesomycetozoea: A Heterogeneous Group of Microorganisms at the Animal-Fungal Boundary. Annu. Rev. Microbiol. 56, 315–344. https://doi.org/10.1146/annurev.micro.56.012302.160950

Mikhailov, K.V., Konstantinova, A.V., Nikitin, M.A., Troshin, P.V., Rusin, L.Y., Lyubetsky, V.A., Panchin, Y.V., Mylnikov, A.P., Moroz, L.L., Kumar, S., Aleoshin, V.V., 2009. The origin of Metazoa: A transition from temporal to spatial cell differentiation. BioEssays 31, 758–768. https://doi.org/10.1002/bies.200800214

Miller, D.J., Hemmrich, G., Ball, E.E., Hayward, D.C., Khalturin, K., Funayama, N., Agata, K., Bosch, T.C.G., 2007. The innate immune repertoire in Cnidaria - Ancestral complexity and stochastic gene loss. Genome Biol. 8. https://doi.org/10.1186/gb-2007-8-4-r59

Mills, D.B., Canfield, D.E., 2014. Oxygen and animal evolution: Did a rise of atmospheric oxygen “trigger” the origin of animals? BioEssays 36, 1145–1155. https://doi.org/10.1002/bies.201400101

Mills, D.B., Ward, L.M., Jones, C., Sweeten, B., Forth, M., Treusch, A.H., Canfield, D.E., 2014. Oxygen requirements of the earliest animals. Proc. Natl. Acad. Sci. 111, 4168–4172. https://doi.org/10.1073/pnas.1400547111

Moroz, L.L., 2015. Convergent evolution of neural systems in ctenophores. https://doi.org/10.1242/jeb.110692

Moroz, L.L., Kocot, K.M., Citarella, M.R., Dosung, S., Norekian, T.P., Povolotskaya, I.S., Grigorenko, A.P., Dailey, C., Berezikov, E., Buckley, K.M., Ptitsyn, A., Reshetov, D., Mukherjee, K., Moroz, T.P., Bobkova, Y., Yu, F., Kapitonov, V.V., Jurka, J., Bobkov, Y.V., Swore, J.J., Girardo, D.O., Fodor, A., Gusev, F., Sanford, R., Bruders, R., Kittler, E., Mills, C.E., Rast, J.P., Derelle, R., Solovyev, V.V., Kondrashov, F.A., Swalla, B.J., Sweedler, J.V., Rogaev, E.I., Halanych, K.M., Kohn, A.B., 2014. The ctenophore genome and the evolutionary origins of neural systems. Nature 510, 109–114. https://doi.org/10.1038/nature13400

Moroz, L.L., Kohn, A.B., 2016. Independent origins of neurons and synapses: Insights from ctenophores. Philos. Trans. R. Soc. B Biol. Sci. 371. https://doi.org/10.1098/rstb.2015.0041

Moroz, L.L., Kohn, A.B., 2015. Unbiased View of Synaptic and Neuronal Gene Complement in Ctenophores: Are There Pan-neuronal and Pan-synaptic Genes across Metazoa? Integr. Comp. Biol. 55, 1028–1049. https://doi.org/10.1093/icb/icv104

Müller, W.E.G., Wiens, M., Adell, T., Gamulin, V., Schröder, H.C., Müller, I.M., 2004. Bauplan of Urmetazoa: Basis for genetic complexity of Metazoa. Int. Rev. Cytol. 235, 53–92. https://doi.org/10.1016/S0074-7696(04)35002-3

Musser, J.M., Schippers, K.J., Nickel, M., Mizzon, G., Kohn, A.B., Pape, C., Hammel, J.U., Wolf, F., Liang, C., Hernández-Plaza, A., Achim, K., Schieber, N.L., Francis, W.R., Vargas R., S., Kling, S., Renkert, M., Feuda, R., Gaspar, I., Burkhardt, P., Bork, P., Beck, M., Kreshuk, A., Wörheide, G., Huerta-Cepas, J.,

Schwab, Y., Moroz, L.L., Arendt, D., 2019. Profiling cellular diversity in sponges informs animal cell type and nervous system evolution (preprint). Developmental Biology. https://doi.org/10.1101/758276

Mylnikov, A.P., Tikhonenkov, D.V., Karpov, S.A., Wylezich, C., 2019. Microscopical Studies on Ministeria vibrans Tong, 1997 (Filasterea) Highlight the Cytoskeletal Structure of the Common Ancestor of Filasterea, Metazoa and Choanoflagellata. Protist 170, 385–396. https://doi.org/10.1016/j.protis.2019.07.001

Narbonne, G.M., 2005. THE EDIACARA BIOTA: Neoproterozoic Origin of Animals and Their Ecosystems. Annu. Rev. Earth Planet. Sci. 33, 421–442. https://doi.org/10.1146/annurev.earth.33.092203.122519

Nedelcu, A.M., 2019. Independent evolution of complex development in animals and plants: deep homology and lateral gene transfer. Dev. Genes Evol. 229, 25–34. https://doi.org/10.1007/s00427-019-00626-8

Nettersheim, B.J., Brocks, J.J., Schwelm, A., Hope, J.M., Not, F., Lomas, M., Schmidt, C., Schiebel, R., Nowack, E.C.M., De Deckker, P., Pawlowski, J., Bowser, S.S., Bobrovskiy, I., Zonneveld, K., Kucera, M., Stuhr, M., Hallmann, C., 2019. Putative sponge biomarkers in unicellular Rhizaria question an early rise of animals. Nat. Ecol. Evol. 3. https://doi.org/10.1038/s41559-019-0806-5

Nichols, S.A., Dirks, W., Pearse, J.S., King, N., 2006. Early evolution of animal cell signaling and adhesion genes. Proc. Natl. Acad. Sci. 103, 12451–12456. https://doi.org/10.1073/pnas.0604065103

Nichols, S.A., Roberts, B.W., Richter, D.J., Fairclough, S.R., King, N., 2012. Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/B-catenin complex. Proc. Natl. Acad. Sci. 109, 13046–13051. https://doi.org/10.1073/pnas.1120685109

Nilsen, T.W., Graveley, B.R., 2010. Expansion of the eukaryotic proteome by alternative splicing. Nature 463, 457–463. https://doi.org/10.1038/nature08909

Nitsche, F., Carr, M., Arndt, H., Leadbeater, B.S.C., 2011. Higher Level Taxonomy and Molecular Phylogenetics of the Choanoflagellatea: CHOANOFLAGELLATE PHYLOGENY. J. Eukaryot. Microbiol. 58, 452–462. https://doi.org/10.1111/j.1550-7408.2011.00572.x

Nitsche, F., Weitere, M., Scheckenbach, F., Hausmann, K., Arndt, H., 2007. Deep Sea Records of Choanoflagellates with a Description of Two New Species 9.

Nyholm, S.V., McFall-Ngai, M., 2004. The winnowing: establishing the squid–vibrio

symbiosis. Nat. Rev. Microbiol. 2, 632–642. https://doi.org/10.1038/nrmicro957

O'Malley, M.A., Wideman, J.G., Ruiz-Trillo, I., 2016. Losing Complexity: The Role of Simplification in Macroevolution. Trends Ecol. Evol. 31, 608–621. https://doi.org/10.1016/j.tree.2016.04.004

Ondracka, A., Dudin, O., Ruiz-Trillo, I., 2018. Decoupling of Nuclear Division Cycles and Cell Size during the Coenocytic Growth of the Ichthyosporean Sphaeroforma arctica. Curr. Biol. 28, 1964-1969.e2. https://doi.org/10.1016/j.cub.2018.04.074

Owczarzak, A., Stibbs, H.H., Bayne, C.J., 1980. The destruction of Schistosoma mansoni mother sporocysts in vitro by amoebae isolated from Biomphalaria glabrata: an ultrastructural study. J. Invertebr. Pathol. 35, 26–33. https://doi.org/10.1016/0022-2011(80)90079-8

Paps, J., Holland, P.W.H., 2018. Reconstruction of the ancestral metazoan genome reveals an increase in genomic novelty. Nat. Commun. 9, 1730. https://doi.org/10.1038/s41467-018-04136-5

Parra-Acero, H., Ros-Rocher, N., Perez-Posada, A., Kożyczkowska, A., Sánchez-Pons, N., Nakata, A., Suga, H., Najle, S.R., Ruiz-Trillo, I., 2018. Transfection of *Capsaspora owczarzaki* , a close unicellular relative of animals. Development 145, dev162107. https://doi.org/10.1242/dev.162107

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S.S., Cepicka, I., Decelle, J., Dunthorn, M., Fiore-Donno, A.M., Gile, G.H., Holzmann, M., Jahn, R., Jirků, M., Keeling, P.J., Kostka, M., Kudryavtsev, A., Lara, E., Lukeš, J., Mann, D.G., Mitchell, E.A.D., Nitsche, F., Romeralo, M., Saunders, G.W., Simpson, A.G.B., Smirnov, A.V., Spouge, J.L., Stern, R.F., Stoeck, T., Zimmermann, J., Schindel, D., de Vargas, C., 2012. CBOL Protist Working Group: Barcoding Eukaryotic Richness beyond the Animal, Plant, and Fungal Kingdoms. PLoS Biol. 10, e1001419. https://doi.org/10.1371/journal.pbio.1001419

Pawlowski, J., Christen, R., Lecroq, B., Bachar, D., Shahbazkia, H.R., Amaral-Zettler, L., Guillou, L., 2011. Eukaryotic Richness in the Abyss: Insights from Pyrotag Sequencing. PLoS ONE 6, e18169. https://doi.org/10.1371/journal.pone.0018169

Pawlowski, J., Lejzerowicz, F., Apotheloz-Perret-Gentil, L., Visco, J., Esling, P., 2016. Protist metabarcoding and environmental biomonitoring: Time for change. Eur. J. Protistol. 55, 12–25. https://doi.org/10.1016/j.ejop.2016.02.003

Perez-Posada, A., Dudin, O., Ocaña-Pallarès, E., Ruiz-Trillo, I., Ondracka, A., 2019. Gradual evolution of cell cycle regulation by cyclin-dependent kinases during the transition to animal multicellularity. bioRxiv 719534. https://doi.org/10.1101/719534

Peterson, K.J., Lyons, J.B., Nowak, K.S., Takacs, C.M., Wargo, M.J., McPeek, M.A., 2004. Estimating metazoan divergence times with a molecular clock. Proc. Natl. Acad. Sci. 101, 6536–6541. https://doi.org/10.1073/pnas.0401670101

Pisani, D., Pett, W., Dohrmann, M., Feuda, R., Rota-Stabelli, O., Philippe, H., Lartillot, N., Wörheide, G., 2015. Genomic data do not support comb jellies as the sister group to all other animals. Proc. Natl. Acad. Sci. 112, 15402–15407. https://doi.org/10.1073/pnas.1518127112

Putnam, N.H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V.V., Jurka, J., Genikhovich, G., Grigoriev, I.V., Lucas, S.M., Steele, R.E., Finnerty, J.R., Technau, U., Martindale, M.Q., Rokhsar, D.S., 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science 317, 86–94. https://doi.org/10.1126/science.1139158

Raghu-kumar, S., 1987. Occurrence of the Thraustochytrid, Corallochytrium limacisporum gen. et sp. nov. in the Coral Reef Lagoons of the Lakshadweep Islands in the Arabian Sea. https://doi.org/10.1515/botm.1987.30.1.83

Richter, D.J., Fozouni, P., Eisen, M.B., King, N., 2018. Gene family innovation, conservation and loss on the animal stem lineage. eLife 7, e34226. https://doi.org/10.7554/eLife.34226

Richter, D.J., King, N., 2013. The Genomic and Cellular Foundations of Animal Origins. Annu. Rev. Genet. 47, 509–537. https://doi.org/10.1146/annurev-genet-111212-133456

Riesgo, A., Farrar, N., Windsor, P.J., Giribet, G., Leys, S.P., 2014. The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. Mol. Biol. Evol. 31, 1102–1120. https://doi.org/10.1093/molbev/msu057

ROBERTS, A., 1969. Conducted Impulses in the Skin of Young Tadpoles. Nature 222, 1265–1266. https://doi.org/10.1038/2221265a0

Robertson, A.D.J., Grutsch, J.F., 1981. Aggregation in Dictyostelium discoideum. https://doi.org/10.1016/0092-8674(81)90087-8

Rogozin, I.B., Carmel, L., Csuros, M., Koonin, E.V., 2012. Origin and evolution of

spliceosomal introns. Biol. Direct 7, 11. https://doi.org/10.1186/1745-6150-7-11

Rokas, A., 2008. The Origins of Multicellularity and the Early History of the Genetic Toolkit For Animal Development. Annu. Rev. Genet. 42, 235–251. https://doi.org/10.1146/annurev.genet.42.110807.091513

Ruiz-Trillo, I., Burger, G., Holland, P.W.H., King, N., Lang, B.F., Roger, A.J., Gray, M.W., 2007. The origins of multicellularity: a multi-taxon genome initiative. Trends Genet. 23, 113–118. https://doi.org/10.1016/j.tig.2007.01.005

Ruiz-Trillo, I., Inagaki, Y., Davis, L.A., Sperstad, S., Landfald, B., Roger, A.J., 2004. Capsaspora owczarzaki is an independent opisthokont lineage. Curr. Biol. 14, R946–R947. https://doi.org/10.1016/j.cub.2004.10.037

Ruiz-Trillo, I., Roger, A.J., Burger, G., Gray, M.W., Lang, B.F., 2008. A Phylogenomic Investigation into the Origin of Metazoa. Mol. Biol. Evol. 25, 664–672. https://doi.org/10.1093/molbev/msn006

Runnegar, B., 1982. Oxygen requirements, biology and phylogenetic significance of the late Precambrian worm Dickinsonia, and the evolution of the burrowing habit. Alcheringa Australas. J. Palaeontol. 6, 223–239. https://doi.org/10.1080/03115518208565415

Ryan, J.F., Pang, K., Schnitzler, C.E., Nguyen, A.D., Moreland, R.T., Simmons, D.K., Koch, B.J., Francis, W.R., Havlak, P., Smith, S.A., Putnam, N.H., Haddock, S.H.D., Dunn, C.W., Wolfsberg, T.G., C.Mullikin, J., Martindale, M.Q., Baxevanis, A.D., 2013. The genome of the ctenophore Mnemiopsis leidyi and its implications for cell type evolution. Science 342. https://doi.org/10.1126/science.1242592

Savin, M.C., Martin, J.L., LeGresley, M., Giewat, M., Rooney-Varga, J., 2004. Plankton diversity in the Bay of Fundy as measured by morphological and molecular methods. Microb. Ecol. 48, 51–65. https://doi.org/10.1007/s00248-003-1033-8

Schaap, P., 2011. Evolutionary crossroads in developmental biology: Dictyostelium discoideum. Development 138, 387–396. https://doi.org/10.1242/dev.048934

Schenkelaars, Q., Pratlong, M., Kodjabachian, L., Fierro-Constain, L., Vacelet, J., Le Bivic, A., Renard, E., Borchiellini, C., 2017. Animal multicellularity and polarity without Wnt signaling. Sci. Rep. 7, 15383. https://doi.org/10.1038/s41598-017-15557-5

Schultheiss, K.P., Craddock, B.P., Suga, H., Miller, W.T., 2014. Regulation of Src and Csk nonreceptor tyrosine kinases in the filasterean Ministeria vibrans. Biochemistry 53, 1320–1329. https://doi.org/10.1021/bi4016499

Schultheiss, K.P., Suga, H., Ruiz-Trillo, I., Miller, W.T., 2012. Lack of Csk-mediated negative regulation in a unicellular SRC kinase. Biochemistry 51, 8267–8277. https://doi.org/10.1021/bi300965h

Sebé-Pedrós, A., Ariza-Cosano, A., Weirauch, M.T., Leininger, S., Yang, A., Torruella, G., Adamski, M., Adamska, M., Hughes, T.R., Gomez-Skarmeta, J.L., Ruiz-Trillo, I., 2013a. Early evolution of the T-box transcription factor family. Proc. Natl. Acad. Sci. 110, 16050–16055. https://doi.org/10.1073/pnas.1309748110

Sebé-Pedrós, A., Ballaré, C., Parra-Acero, H., Chiva, C., Tena, J.J., Sabidó, E., Gómez-Skarmeta, J.L., Di Croce, L., Ruiz-Trillo, I., 2016a. The Dynamic Regulatory Genome of Capsaspora and the Origin of Animal Multicellularity. Cell 165, 1224–1237. https://doi.org/10.1016/j.cell.2016.03.034

Sebé-Pedrós, A., Chomsky, E., Pang, K., Lara-Astiaso, D., Gaiti, F., Mukamel, Z., Amit, I., Hejnol, A., Degnan, B.M., Tanay, A., 2018a. Early metazoan cell type diversity and the evolution of multicellular gene regulation. Nat. Ecol. Evol. 2, 1176–1188. https://doi.org/10.1038/s41559-018-0575-6

Sebé-Pedrós, A., de Mendoza, A., Lang, B.F., Degnan, B.M., Ruiz-Trillo, I., 2011. Unexpected Repertoire of Metazoan Transcription Factors in the Unicellular Holozoan Capsaspora owczarzaki. Mol. Biol. Evol. 28, 1241–1254. https://doi.org/10.1093/molbev/msq309

Sebé-Pedrós, A., Degnan, B.M., Ruiz-Trillo, I., 2017. The origin of Metazoa: a unicellular perspective. Nat. Rev. Genet. 18, 498–512. https://doi.org/10.1038/nrg.2017.21

Sebé-Pedrós, A., Irimia, M., del Campo, J., Parra-Acero, H., Russ, C., Nusbaum, C., Blencowe, B.J., Ruiz-Trillo, I., 2013b. Regulated aggregative multicellularity in a close unicellular relative of metazoa. eLife 2, e01287. https://doi.org/10.7554/eLife.01287

Sebé-Pedrós, A., Peña, M.I., Capella-Gutiérrez, S., Antó, M., Gabaldón, T., Ruiz-Trillo, I., Sabidó, E., 2016b. High-Throughput Proteomics Reveals the Unicellular Roots of Animal Phosphosignaling and Cell Differentiation. Dev. Cell 39, 186–197. https://doi.org/10.1016/j.devcel.2016.09.019

Sebé-Pedrós, A., Roger, A.J., Lang, F.B., King, N., Ruiz-Trillo, I., 2010. Ancient origin of the integrin-mediated adhesion and signaling machinery. Proc. Natl. Acad. Sci. 107, 10142–10147. https://doi.org/10.1073/pnas.1002257107

Sebé-Pedrós, A., Saudemont, B., Chomsky, E., Plessier, F., Mailhé, M.-P., Renno, J., Loe-Mie, Y., Lifshitz, A., Mukamel, Z., Schmutz, S., Novault, S., Steinmetz,

P.R.H., Spitz, F., Tanay, A., Marlow, H., 2018b. Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. Cell 173, 1520-1534.e20. https://doi.org/10.1016/j.cell.2018.05.019

Sebé-Pedrós, A., Zheng, Y., Ruiz-Trillo, I., Pan, D., 2012. Premetazoan origin of the hippo signaling pathway. Cell Rep. 1, 13–20. https://doi.org/10.1016/j.celrep.2011.11.004

Shalchian-Tabrizi, K., Minge, M.A., Espelund, M., Orr, R., Ruden, T., Jakobsen, K.S., Cavalier-Smith, T., 2008. Multigene Phylogeny of Choanozoa and the Origin of Animals. PLoS ONE 3, e2098. https://doi.org/10.1371/journal.pone.0002098

Siebert, S., Farrell, J.A., Cazet, J.F., Abeykoon, Y., Primack, A.S., Schnitzler, C.E., Juliano, C.E., 2019. Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution. Science 365, eaav9314. https://doi.org/10.1126/science.aav9314

Simakov, O., Kawashima, T., 2017. Independent evolution of genomic characters during major metazoan transitions. Dev. Biol. 427, 179–192. https://doi.org/10.1016/j.ydbio.2016.11.012

Sogabe, S., Hatleberg, W.L., Kocot, K.M., Say, T.E., Stoupin, D., Roper, K.E., Fernandez-Valverde, S.L., Degnan, S.M., Degnan, B.M., 2019. Pluripotency and the origin of animal multicellularity. Nature 570, 519–522. https://doi.org/10.1038/s41586-019-1290-4

Sperling, E.A., Frieder, C.A., Raman, A.V., Girguis, P.R., Levin, L.A., Knoll, A.H., 2013. Oxygen, ecology, and the Cambrian radiation of animals. Proc. Natl. Acad. Sci. 110, 13446–13451. https://doi.org/10.1073/pnas.1312778110

Srivastava, M., Begovic, E., Chapman, J., Putnam, N.H., Hellsten, U., Kawashima, T., Kuo, A., Mitros, T., Salamov, A., Carpenter, M.L., Signorovitch, A.Y., Moreno, M.A., Kamm, K., Grimwood, J., Schmutz, J., Shapiro, H., Grigoriev, I.V., Buss, L.W., Schierwater, B., Dellaporta, S.L., Rokhsar, D.S., 2008. The Trichoplax genome and the nature of placozoans. Nature 454, 955–960. https://doi.org/10.1038/nature07191

Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M.E.A., Mitros, T., Richards, G.S., Conaco, C., Dacre, M., Hellsten, U., Larroux, C., Putnam, N.H., Stanke, M., Adamska, M., Darling, A., Degnan, S.M., Oakley, T.H., Plachetzki, D.C., Zhai, Y., Adamski, M., Calcino, A., Cummins, S.F., Goodstein, D.M., Harris, C., Jackson, D.J., Leys, S.P., Shu, S., Woodcroft, B.J., Vervoort, M., Kosik, K.S., Manning, G., Degnan, B.M., Rokhsar, D.S., 2010. The

Amphimedon queenslandica genome and the evolution of animal complexity. Nature 466, 720–6. https://doi.org/10.1038/nature09201

Stanley, S.M., 1973. An Ecological Theory for the Sudden Origin of Multicellular Life in the Late Precambrian. Proc. Natl. Acad. Sci. 70, 1486–1489. https://doi.org/10.1073/pnas.70.5.1486

Steenkamp, E.T., Wright, J., Baldauf, S.L., 2006. The Protistan Origins of Animals and Fungi. Mol. Biol. Evol. 23, 93–106. https://doi.org/10.1093/molbev/msj011

Steinmetz, P.R.H., Kraus, J.E.M., Larroux, C., Hammel, J.U., Amon-Hassenzahl, A., Houliston, E., Wörheide, G., Nickel, M., Degnan, B.M., Technau, U., 2012. Independent evolution of striated muscles in cnidarians and bilaterians. Nature 487, 231–234. https://doi.org/10.1038/nature11180

Stibbs, H.H., Owczarzak, A., Bayne, C.J., DeWan, P., 1979. Schistosome sporocyst-killing amoebae isolated from Biomphalaria glabrata. J. Invertebr. Pathol. 33, 159–170. https://doi.org/10.1016/0022-2011(79)90149-6

Stougaard, P., Jørgensen, F., Johnsen, M.G., Hansen, O.C., 2002. Microbial diversity in ikaite tufa columns: an alkaline, cold ecological niche in Greenland. Environ. Microbiol. 4, 487–493.

Suga, H., Chen, Z., de Mendoza, A., Sebé-Pedrós, A., Brown, M.W., Kramer, E., Carr, M., Kerner, P., Vervoort, M., Sánchez-Pons, N., Torruella, G., Derelle, R., Manning, G., Lang, B.F., Russ, C., Haas, B.J., Roger, A.J., Nusbaum, C., Ruiz-Trillo, I., 2013. The Capsaspora genome reveals a complex unicellular prehistory of animals. Nat. Commun. 4, 2325. https://doi.org/10.1038/ncomms3325

Suga, H., Dacre, M., de Mendoza, A., Shalchian-Tabrizi, K., Manning, G., Ruiz-Trillo, I., 2012. Genomic Survey of Premetazoans Shows Deep Conservation of Cytoplasmic Tyrosine Kinases and Multiple Radiations of Receptor Tyrosine Kinases. Sci. Signal. 5, ra35–ra35. https://doi.org/10.1126/scisignal.2002733

Suga, H., Miller, W.T., 2018. Src signaling in a low-complexity unicellular kinome. Sci. Rep. 8, 5362. https://doi.org/10.1038/s41598-018-23721-8

Suga, H., Ruiz-Trillo, I., 2013. Development of ichthyosporeans sheds light on the origin of metazoan multicellularity. Dev. Biol. 377, 284–292. https://doi.org/10.1016/j.ydbio.2013.01.009

Suga, H., Torruella, G., Burger, G., Brown, M.W., Ruiz-Trillo, I., 2014. Earliest Holozoan expansion of phosphotyrosine signaling. Mol. Biol. Evol. 31, 517–528. https://doi.org/10.1093/molbev/mst241

Szathmáry, E., Smith, J.M., 1995. The major evolutionary transitions. Nature 374, 227–232. https://doi.org/10.1038/374227a0

Takishita, K., Miyake, H., Kawato, M., Maruyama, T., 2005. Genetic diversity of microbial eukaryotes in anoxic sediment around fumaroles on a submarine caldera floor based on the small-subunit rDNA phylogeny. Extrem. Life Extreme Cond. 9, 185–196. https://doi.org/10.1007/s00792-005-0432-9

Takishita, K., Tsuchiya, M., Kawato, M., Oguri, K., Kitazato, H., Maruyama, T., 2007. Genetic Diversity of Microbial Eukaryotes in Anoxic Sediment of the Saline Meromictic Lake Namako-ike (Japan): On the Detection of Anaerobic or Anoxic-tolerant Lineages of Eukaryotes. Protist 158, 51–64. https://doi.org/10.1016/j.protis.2006.07.003

Takishita, K., Tsuchiya, M., Reimer, J.D., Maruyama, T., 2006. Molecular evidence demonstrating the basidiomycetous fungus Cryptococcus curvatus is the dominant microbial eukaryote in sediment at the Kuroshima Knoll methane seep. Extrem. Life Extreme Cond. 10, 165–169. https://doi.org/10.1007/s00792-005-0495-7

Telford, M.J., Budd, G.E., Philippe, H., 2015. Phylogenomic Insights into Animal Evolution. Curr. Biol. 25, R876–R887. https://doi.org/10.1016/j.cub.2015.07.060

Thomsen, H.A., Nitsche, F., Richter, D.J., 2016. Seasonal Occurrence of Loricate Choanoflagellates in Danish Inner Waters. Protist 167, 622–638. https://doi.org/10.1016/j.protis.2016.09.002

Tong, S.M., Nygaard, K., Bernard, C., Vørs, N., Patterson, D.J., 1998. Heterotrophic flagellates from the water column in Port Jackson, Sydney, Australia. Eur. J. Protistol. 34, 162–194. https://doi.org/10.1016/S0932-4739(98)80027-8

Torruella, G., de Mendoza, A., Grau-Bové, X., Antó, M., Chaplin, M.A., del Campo, J., Eme, L., Pérez-Cordón, G., Whipps, C.M., Nichols, K.M., Paley, R., Roger, A.J., Sitjà-Bobadilla, A., Donachie, S., Ruiz-Trillo, I., 2015. Phylogenomics Reveals Convergent Evolution of Lifestyles in Close Relatives of Animals and Fungi. Curr. Biol. 25, 2404–2410. https://doi.org/10.1016/j.cub.2015.07.053

Torruella, G., Derelle, R., Paps, J., Lang, B.F., Roger, A.J., Shalchian-Tabrizi, K., Ruiz-Trillo, I., 2012. Phylogenetic Relationships within the Opisthokonta Based on Phylogenomic Analyses of Conserved Single-Copy Protein Domains. Mol. Biol. Evol. 29, 531–544. https://doi.org/10.1093/molbev/msr185

Vinayak, S., Pawlowic, M.C., Sateriale, A., Brooks, C.F., Studstill, C.J., Bar-Peled, Y., Cipriano, M.J., Striepen, B., 2015. Genetic modification of the diarrhoeal

pathogen Cryptosporidium parvum. Nature 523, 477–480. https://doi.org/10.1038/nature14651

Wainright, P.O., Hinkle, G., Sogin, M.L., Stickel, S.K., 1993. Monophyletic origins of the metazoa: an evolutionary link with fungi. Science 260, 340–342. https://doi.org/10.1126/science.8469985

Wetzel, L.A., Levin, T.C., Hulett, R.E., Chan, D., King, G.A., Aldayafleh, R., Booth, D.S., Sigg, M.A., King, N., 2018. Predicted glycosyltransferases promote development and prevent spurious cell clumping in the choanoflagellate S. rosetta. eLife 7, e41482. https://doi.org/10.7554/eLife.41482

Whelan, N.V., Kocot, K.M., Moroz, L.L., Halanych, K.M., 2015. Error, signal, and the placement of Ctenophora sister to all other animals. Proc. Natl. Acad. Sci. 112, 5773–5778. https://doi.org/10.1073/pnas.1503453112

Windsor Reid, P.J., Matveev, E., McClymont, A., Posfai, D., Hill, A.L., Leys, S.P., 2018. Wnt signaling and polarity in freshwater sponges. BMC Evol. Biol. 18, 12. https://doi.org/10.1186/s12862-018-1118-0

Worley, A.C., Raper, K.B., Hohl, M., 1979. Fonticula alba: A New Cellular Slime Mold (Acrasiomycetes). Mycologia 71, 746–760. https://doi.org/10.2307/3759186

Woznica, A., Cantley, A.M., Beemelmanns, C., Freinkman, E., Clardy, J., King, N., 2016. Bacterial lipids activate, synergize, and inhibit a developmental switch in choanoflagellates. Proc. Natl. Acad. Sci. 113, 7894–7899. https://doi.org/10.1073/pnas.1605015113

Woznica, A., Gerdt, J.P., Hulett, R.E., Clardy, J., King, N., 2017. Mating in the Closest Living Relatives of Animals Is Induced by a Bacterial Chondroitinase. Cell 170, 1175-1183.e11. https://doi.org/10.1016/j.cell.2017.08.005

Woznica, A., King, N., 2018. Lessons from simple marine models on the bacterial regulation of eukaryotic development. Curr. Opin. Microbiol. 43, 108–116. https://doi.org/10.1016/j.mib.2017.12.013

Wylezich, C., Karpov, S.A., Mylnikov, A.P., Anderson, R., Jürgens, K., 2012. Ecologically relevant choanoflagellates collected from hypoxic water masses of the Baltic Sea have untypical mitochondrial cristae. BMC Microbiol. 12, 271. https://doi.org/10.1186/1471-2180-12-271

Xiao, S., Knoll, A.H., Schiffbauer, J.D., Zhou, C., Yuan, X., 2012. Comment on “Fossilized nuclei and germination structures identify Ediacaran ‘animal embryos’ as encysting protists.” Science 335, 1169; author reply 1169. https://doi.org/10.1126/science.1218814

Young, S.L., Diolaiti, D., Conacci-Sorrell, M., Ruiz-Trillo, I., Eisenman, R.N., King, N., 2011. Premetazoan Ancestry of the Myc–Max Network. Mol. Biol. Evol. 28, 2961–2971. https://doi.org/10.1093/molbev/msr132

Zakhvatkin AA, 1949. The comparative embryology of the low invertebrates., in: Sources and Method of the Origin of Metazoan Development. Soviet Science.

Zettler, L.A.A., Nerad, T.A., O'Kelly, C.J., Sogin, M.L., 2001. The Nucleariid Amoebae: More Protists at the Animal-Fungal Boundary. J. Eukaryot. Microbiol. 48, 293–297. https://doi.org/10.1111/j.1550-7408.2001.tb00317

4. Discussion

Discussion

The work presented in this thesis contributes to the development of *Capsaspora owczarzaki* as a model system, by characterizing its cell cycle regulation (Result 3.1) and by establishing experimental protocols such as transgenesis (Result 3.2) and culture synchronization (Result 3.1). This chapter is structured in two sections, the first deepening and discussing the cell cycle of Capsaspora, and the second framing the development of tools for Capsaspora under the context of model organisms to study the origin of animal multicellularity.

4.1 The cell cycle regulation of *Capsaspora*

In this study, we have developed a method of synchronization of *Capsaspora* adherent cell cultures and illustrate its applicability by performing a transcriptomic characterization of the cell cycle in an organism beyond the conventional catalogue of model species. This approach has allowed us to gain insight into key aspects of the cell cycle -such as cell division and periodic gene expression- in a unicellular relative of animals. Our study of gene expression during the cell cycle has shown that transcriptional regulation occurs concomitant to the progression through the cell cycle in *Capsaspora*. Likewise, it has allowed establishing more connections between many *Capsaspora* genes and their known cell cycle roles in other organisms.

a) Uniqueness and similarities of the *Capsaspora* cell cycle

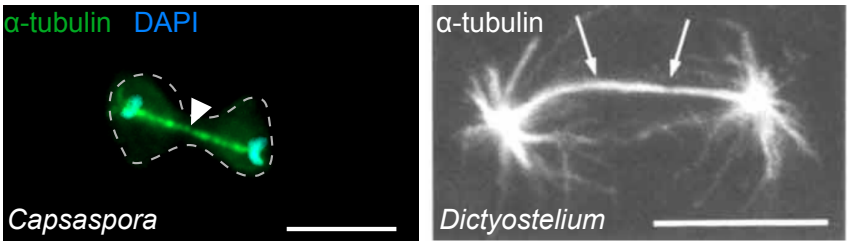
Capsaspora cell cycle regulation seems to resemble a more ancestral state than yeasts or humans, illustrated by its genetic repertoire of the cyclin-CDK system, the presence of E2F and Rb genes, and the presence of orthologs present in outgroup lineages outside Opisthokonta (such as putative components of the RZZ complex in *Dictyostelium discoideum* (Samerier, 2011)). While we observed a fairly conserved machinery of cell cycle entry, DNA replication, and mitosis, it is interesting to note the absence of typical animal- or yeast-like CDK inhibitors in *Capsaspora*, such as p16 or p21, which control entry into the cell cycle. Thus, we wonder about the nature of the regulation to enter the cell cycle in this species. It is very likely that *Capsaspora* is able to monitor intra- and extracellular cues in a similar fashion to other unicellular organisms like yeasts, by

sensing changes in the cytoplasm/nucleus ratio and/or indirectly sensing nutrient abundance, although it was not under the scope of our study to identify such candidate regulators of this process.

Based on our observations of gene expression and transcription factor motifs (see Appendix 1), we hypothesize that *Capsaspora* may have a putative network of transcription factors controlling the entry into the cell cycle, where E2F7 represses E2F1 transcription while Rb inhibits E2F1 activity. Upon entry into cell cycle mediated by MAP kinases, these inhibitors may disassemble and allow the activation of E2F1, which would initialize transcription of G1/S genes. Among these genes, some are transcription factors, such as b-Myb, that may regulate gene expression at further stages of the cell cycle. Although our current data to support these hypotheses are mere computational predictions, this suggests that control of cell cycle gene expression by E2F + b-Myb might be conserved in *Capsaspora*, as happens in other species (Fischer and Müller, 2017). Cell cycle regulation is also driven by ubiquitination-mediated protein degradation; however, only a small number of ubiquitin ligase complex subunits are periodically regulated in *Capsaspora*. From all the subunit of the Anaphase-Promoter Complex (APC) currently described in unicellular relatives of animals (Eme et al., 2011), only UBE2S shows periodic activity. Similarly, only the SCF complex subunit BTRC (another ubiquitin-ligase complex that regulates the APC and progression through the cell cycle (Morgan, 2007)) shows a peak of expression during the cell cycle. One possible interpretation is that regulation of these genes occurs at the protein level and not at the transcriptional level, as also suggested for the E2F1 transcription factor and other cell cycle regulators.

Thus, the regulation of the cell cycle in *Capsaspora* shares many commonalities with other eukaryotes. However, cell cycle regulation can be deployed very differently across species (Jensen et al., 2006). This is reflected by the gene age of the *Capsaspora* periodic transcriptional program. Despite the otherwise conserved program of mitotic cell division as shown by individual inspection of genes, the G2/M and M gene clusters do not appear as significantly enriched in any particular gene age, in either of the iterations analysed. One way to interpret this is that the *Capsaspora* mitotic gene expression program contains as many old, typical eukaryotic genes as newly innovated, or even *Capsaspora*-specific genes. This contrasts with the largely conserved gene set of DNA replication and entry into cell cycle, and may be related to specificities in the mechanisms of cell division in *Capsaspora*, such as its fast cytokinesis. Cell division in *Capsaspora*

shows traits of a typical amoeboid cell; microtubule structures bear a high resemblance to previous reports of mitotic cell division in *D. discoideum* amoebae (Fukui and Inoue, 1991), with a bundle of bridging microtubules prominently visible (Fig. 13). Whereas conservation of the regulators and machinery of the spindle and mitotic checkpoints is widely conserved in *Capsaspora*, the absence of typical centrosomal or spindle pole body proteins in this organism raises new questions about how microtubules assemble and arrange to form the spindle in this species. As discussed in Result 3.1, two possibilities are that *Capsaspora* microtubules are able to self-assemble, or that *Capsaspora* relies on a microtubule-organizer machinery independently acquired at some point in the lineage that led to *Capsaspora*. Further investigation could focus in the isolation of gamma-tubulin-associated proteins in Capsaspora, either by pull-down or co-immunoprecipitation, as gamma-tubulin acts as a nucleator to polymerize microtubules at the microtubule-organizing center.



Perez-Posada et al., 2019 Fukui and Inoue, 1991

Fig. 13: The structure of the mitotic spindle in *Capsaspora* resembles that of *Dictyostelium discoideum*. Arrows indicate the bridging microtubules of the mitotic spindle. Adapted from Perez-Posada et al., 2019 and Fukui and Inoue, 1991.

As an unicellular organism, it is likely that the cell cycle of Capsaspora is tightly related to its life cycle. For example, cell cycle might be putatively down-regulated upon entering in the cystic stage, matching previous observations during its life cycle (Sebé-Pedrós et al., 2013). It is well known that Capsaspora can form aggregates of independent cells coming together. Recent observations of an aggregate-triggering factor in the culture medium of *Capsaspora* (Ros-Rocher, Gerdt, personal communication) and the putative presence of cell types within *Capsaspora* aggregates (Najle et al., personal communication) suggest that the biology of *Capsaspora* aggregates might be more complex than previously stated. Although not used as a layer to regulate controlled proliferation in a multicellular bauplan, I speculate that *Capsaspora* cell cycle phases could be able to determine cell identity within the aggregate (only if proven that different cell types exist within aggregates), following prior findings in other aggregative species like *D. discoideum* (Gomer and Firtel, 1987; Soufi and Dalton, 2016).

b) The cyclin-CDK system of Holozoa in the context of eukaryotic evolution

The cyclin-CDK repertoire of Capsaspora is more similar to other opisthokonts outside yeasts, its genome encoding cyclins from the A, B and E families. We detected a conserved pattern of expression of cyclins E, A, B, and CDK1, between Capsaspora and human cells. We think that *Capsaspora* cyclins E, A and B may bind sequentially to CDK1 throughout the cell cycle, in a similar situation to *S. cerevisiae* cell cycle and likely to other opisthokonts and eukaryotes, but with a cyclin repertoire more similar to that of animals. Future confirmations of these proposed interactions could be done via binding assays of Capsaspora CDK1 with E, A, and B cyclins, including pull-down assays (Draetta et al., 1989), two-hybrid assays (see (Measday et al., 1997) for a former detection of cyclin partners in *S. cerevisiae*), or genetic approaches if available.

That being the case, our observations would support our proposed model for the evolution of the cyclin/CDK system within Opisthokonta (Fig. 14). We consider a scenario matching previous models of ancient eukaryotic cell cycle regulation, where a primitive cyclin-CDK oscillator drives periodic DNA replication and mitosis and additional regulatory layers were implemented throughout the course of evolution in different lineages (Nasmyth, 1995), including independent expansions and sub-functionalizations of cyclins and CDKs. Such regulation is endorsed by systems engineered for a single cyclin/CDK oscillation driving progression through the cell cycle, such as yeast or knockout mice embryos, or by organisms well known to have a single CDK involved in the cell cycle (Coudreuse and Nurse, 2010; Gutierrez, 2009; Murray and Kirschner, 1989; Santamaría et al., 2007).

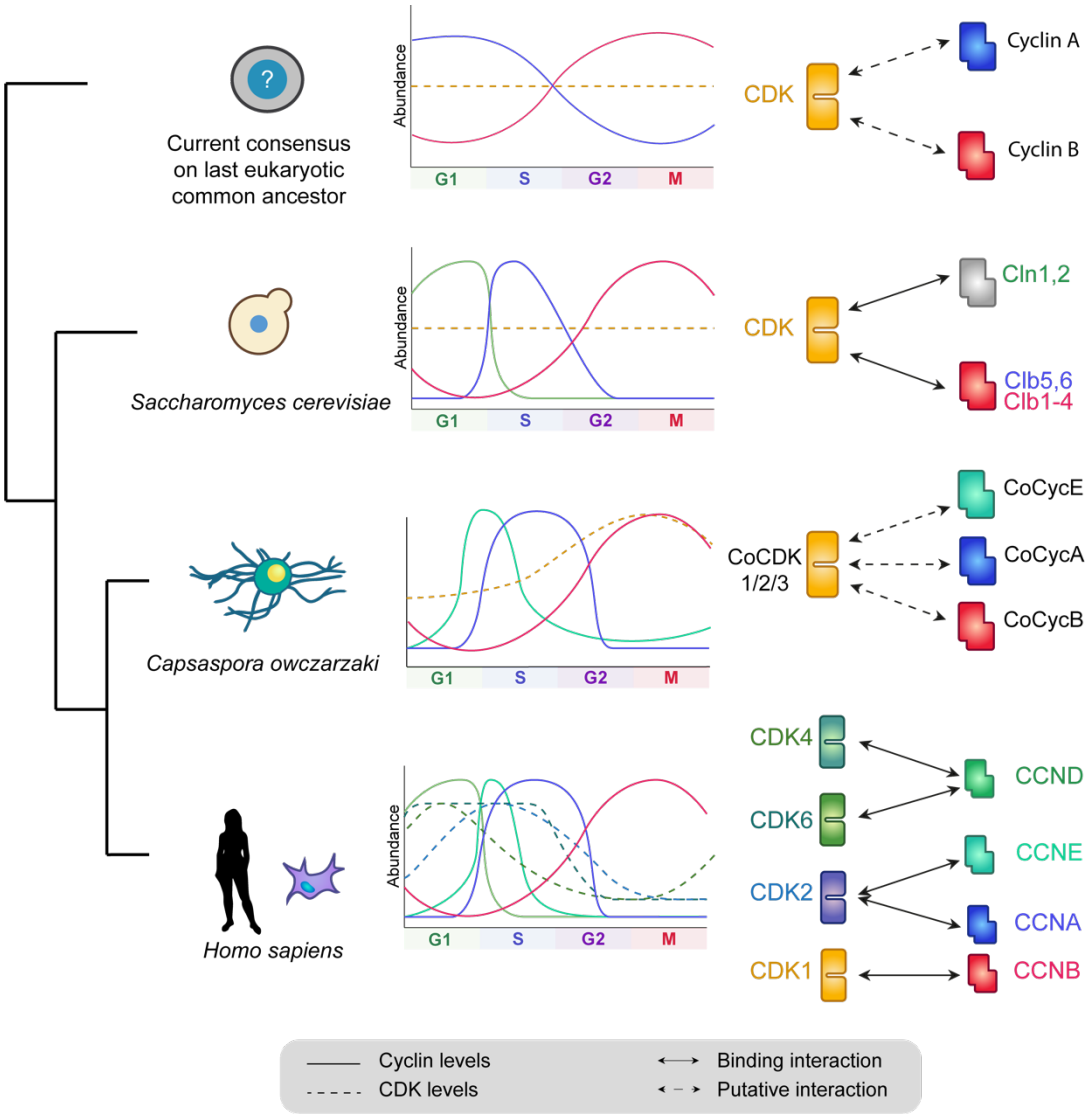


Fig. 14: a model of cell cycle regulation by cyclins and CDKs in Opisthokonta.

Based on phylogenies by multiple authors ((Cao et al., 2014; Medina et al., 2016), ourselves) (see Appendix 1), the last eukaryotic common ancestor (LECA) likely had A- and B-type cyclins, among others including cyclins G/I, but not D- and E-type cyclins (Fig. 15). In Opisthokonta, fungi lost cyclin A, independently expanding cyclin B into Cln- and Clb-type cyclins. Cyclin E would have appeared in the last common ancestor of Holozoa, and cyclin D later in the lineage that led to animals, together with the expansion of CDK1 into CDK2,3,4,6 (Cao et al., 2014) (Fig. 15). This model would imply different expansions of cyclins (mostly B) and CDKs in eukaryotes, with different lineages exhibiting convergent evolution of a G1-phase cyclin responding to extracellular cues (the term “cyclin D” in Archaeplastida being misleadingly referring to a protein that is sister group to a different subfamily of cyclins, but homolog in function to metazoan cyclin

D). The latter would not be surprising, provided the known redundancy in function of cyclins and CDKs and the phylogenetic proximity of metazoan cyclin D to the cyclin G/I family (Medina et al., 2016). It could also be possible that the ancestral function of G/I cyclins was, precisely, to trigger the entry into the cell cycle, provided that Rb and E2F were likely present in the LECA (de Mendoza and Sebé-Pedrós, 2019; Desvoyes et al., 2014). Our model, while matching previous observations in phylogenies and functional studies, is limited due to the difficulty of inferring strong phylogenetic relationships between cyclin and CDK subfamilies, as they show very little conservation outside the cyclin box and PSTAIRE motifs (Morgan, 2007). In addition, for a better understanding of the cyclin-CDK regulation in other eukaryotes and further validation of our model, more functional studies are needed outside the current set of studied species.

Although we focused on *Capsaspora*, I find interesting to discuss the phylogenetic profile of cyclins and CDKs in other unicellular holozoans. We found an internal duplication of both cyclin B and CDK1 within choanoflagellates, of cyclin B in the filasterean *Pigoraptor chileana*, and multiple co-orthologs of cyclin B1-2 in the filasterean *Ministeria vibrans* (Appendix 1). Although *Monosiga brevicollis* and *S. rosetta* do not seem to have a cyclin E ortholog, it cannot be excluded an independent loss and that it may be present in other choanoflagellates. This scattering and independent expansion of the cyclin/CDK repertoire in choanoflagellates and filastereans, known to have life cycles with multiple life stages and cell types, results of great interest for further interrogating the basis of cell cycle regulation in the origin of animals. Such expansions recapitulate the aforementioned expansions of cyclins and CDKS after the last common ancestor of animals and choanoflagellates (Cao et al., 2014), which we also recover in our phylogenies. It becomes interesting then to consider the putative implications of G1 cyclins and specific CDKs in the emergence of more complex cell cycle regulation, adding an additional layer of regulation for additional proliferation/inhibitory signals beyond the environmental, non-cellular cues. Such window of regulation might have contributed to the coordination of cell reproduction in a multicellular context, as previously discussed by (Cao et al., 2014) (see Result 3.3).

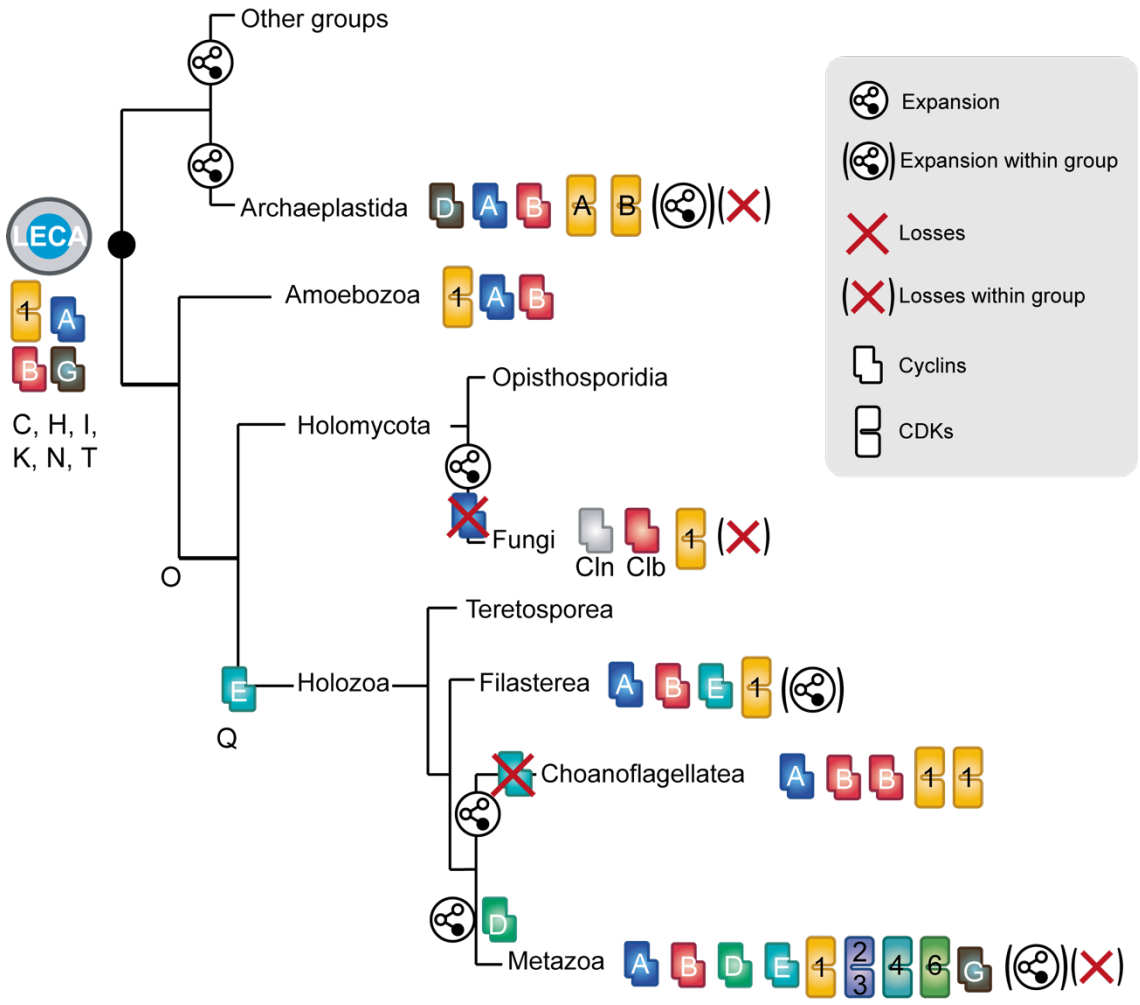


Figure 15: a possible reconstruction of the evolution of the animal cyclin-CDK cell cycle regulator kit. Phylogenetic tree of eukaryotes displaying the different expansions, gains and losses of the cyclin and CDK subfamilies with a known role in cell cycle regulation, in different eukaryotic groups. Events within group refers to some species within the group in which show additional expansions, gains or losses beside those depicted here. Note that LECA cyclin G/I likely gave rise to plants cyclin D and share a common ancestor to the metazoan cyclins D and G/I. Gain of other cyclin families are depicted as regular letters at different stems.

c) Conservation and divergence of the cell cycle regulation

Evolution and conservation of periodic gene expression might not have been a key factor to explain similarities and differences between animals and their unicellular relatives, due to the many regulatory layers responsible for periodic activity during the cell cycle. Regulation of the cell cycle is long-known to occur at multiple levels, from the acquisition of temporal expression to temporal regulation in the form of translational regulation, or protein activation-repression (Dunphy and Newport, 1988). These multiple regulatory layers allow the deployment of cell cycle programs in many different ways across cell types and species, despite a core machinery inherent to the need to replicate

and segregate DNA during cell reproduction. Our comparative analyses of periodic transcriptional programs suggest that pairwise species comparison alone performs poorly as a proxy of how similar is cell cycle regulation between two species compared to a third species. Conclusions drawn from our and similar studies are limited for multiple reasons:

- First, genomic comparisons between distantly related species are usually possible only at orthogroup levels, due to difficulties in inferring orthology relationships due to events of gene loss, expansion, and shuffling, occurring across large evolutionary scales (Tekaia, 2016). Comparisons are even more limited due to our lack of experimental validation beyond sequence homology.
- There are several aspects of cell cycle regulation that are not captured by our current method of analysis and comparison. Gene expression can quickly diverge in evolution by losing or acquiring TF binding sites in cis-regulatory sequences such as promoters or enhancers (Wray, 2007). This leads to divergences in the periodic transcriptional programs across species, which may (or may not) translate into keeping the periodic activity of a complex by recruiting different partners of the same regulatory network in different lineages through different mechanisms (Jensen et al., 2006). In addition, comparisons of the transcriptional regulation during the cell cycle might not reflect the actual similarities and differences as post-transcriptional processes such as splicing or translation are also very dynamically regulated throughout the cell cycle, making mRNA-protein levels correlate differently for each gene (Aviner et al., 2013; Kronja and Orr-Weaver, 2011; Liu et al., 2016; Novoa et al., 2010; Sivan and Elroy-Stein, 2008; Stumpf et al., 2013). This goes hand in hand with the fact that there is no clear distinction between periodic and non-periodic genes only by gene expression, especially for lowly expressed or low-dynamically expressed genes, as fluctuations in transcription can translate very differently for each particular case.
- A substantial source of bias comes from the currently available datasets of different species and cell types. A larger homogeneity in the cell cycle program is expected in datasets from the same unicellular cell type than in datasets from different cell lines; the differences are more evident if some of these cell lines show aberrant gene expression and overall biology as a result of immortalization (See (Ben-David et al., 2018, 2017; Kobayashi et al., 2015b; Landry et al., 2013) for examples aberrant genome, gene expression, and heterogeneity). In addition,

different structures in cell cycle phasing across species may add an extra layer of complexity when interrogating for the conservation of phase-dependent expression (Jensen et al., 2006; Kelliher et al., 2016).

- Finally, the different studies are undoubtedly biased by experimental, technical, and methodological differences: differences in sample treatments, timing and conditions, differences and biases from experimental techniques (microarrays and next-generation sequencing; see (Zhao et al., 2014) for a case example), differences in detection algorithms and thresholds applied, or even bias introduced by missing and obsolete data, such as microarray identifiers.

As a line of research with a strong component of comparative approaches, investigating conservation and evolution of the periodic transcriptional across eukaryotes entails additional challenges because of the aforementioned reasons. Future comparative approaches exploring the evolution of cell cycle regulation in opisthokonts should focus on the following main fronts:

- Including more non-immortalized cell types from more different animals and other multicellular species, in order to understand the differences across cell types within one single organism.
- Exploring cell cycle regulation in other unicellular relatives of animals such as choanoflagellates, ichthyosporeans, or other filastereans (Hehenberger et al., 2017; Torruella et al., 2015).
- Exploring cell cycle gene expression in non-model, non-holozoan opisthokonts such as nucleariids, fonticulids (Brown and Silberman, 2013; López-Escardó et al., 2018; Worley et al., 1979), and other fungal species (Hong et al., 2014; Klar, 2013), to contextualize cell cycle regulation of yeasts as model systems provided their divergence and fast evolution in relation to the rest of fungi.
- Studying other groups within the branch of Amorphea, such as Apusomonadida, and Amoebozoa (Dove et al., 1986), for a better understanding of pre-Opisthokonta PTP and to infer the PTP in the last common opisthokont ancestor.
- Studying the conservation of other regulatory layers of the cell cycle besides periodic transcription. Together with recent advances into the proteomic characterization of *Capsaspora* (Sebé-Pedrós et al., 2016b), our protocol of cell cycle synchronization opens the door for studies of changes in the phosphoproteome of *Capsaspora* during the cell cycle, which could unravel the

post-translational regulation of many genes that might complement our current knowledge of cell cycle regulation in *Capsaspora*.

Assessing the problem of taxon bias has been classically difficult due to the need to isolate and grow cells in cultures in enough quantity and for long enough to retrieve data during a whole cell cycle event; and secondly, due to the inability to synchronize these cell cultures. These hindrances may be overcome in the future by the advent of single-cell technologies, whereby high-throughput gene expression data can be obtained at the independent, cellular level, removing the need to obtain culturable cells and synchronized cultures. These methods are able to retrieve gene expression data relative to the cell cycle relatively easily (Barron and Li, 2016) and abundantly enough to allow for further analysis, allowing the reconstruction of cell cycle pseudo-timelines (Barry et al., 2015; Karlsson et al., 2017; Proserpio et al., 2015; Scialdone et al., 2015), thus having the potential to uncover differences between cell cycle regulation in different cell types within and between organisms.

As an initial step, these tools and approaches may help to better understand the origin of the cell cycle-associated transcriptional regulation found in multicellular organisms. Ultimately, however, functional validation is crucial to infer the role and involvement of periodic genes in each species, complementing our vision of the periodic transcriptional regulation and corroborating our hypothesis regarding the evolution of cell cycle regulations in opisthokonts.

4.2 Development of *Capsaspora* as a model system and the value of novel model organisms

Some questions in biology escape the current landscape of tools available for research. As previously stated, the evolution of the cell cycle is one among many questions limited by the taxon bias, that is, the lack of organisms experimentally tractable outside a selected number of species. This situation, observable across multiple disciplines, has prompted a steady effort into developing tools for multiple non-model organisms throughout eukaryotes, including *Capsaspora* and other unicellular relatives for the question of the transition to animal multicellularity.

The origin of multicellularity, a feature with an undeniable role in the emergence of animals, has its roots in the lesser level of organization shown by its current, extant unicellular relatives. Understanding the mechanisms that facilitated, preceded, and

fostered multicellularity is not possible unless we attempt to understand the nature of two types of organisms key in the transition to animal multicellularity: the last common ancestor of animals, and the unicellular ancestors in the stem that led to the animal lineage (reviewed in section 3.3). Efforts to understand this transition has focused mostly on the reconstruction of the genetic content of the ancestor. These reconstructions have shown that many genes previously thought to be exclusively present in animals were also present in their unicellular ancestor. Some of these genes are known to be crucial in multicellular processes, such as development, and yet they are present in unicellular organisms. The next logical step, therefore, is to investigate the function of these genes in a unicellular context, that is, the biology of the extant unicellular relatives of animals.

In order to investigate the biology of these organisms, we need tools that are able to tell us about the function of their genes. Whereas our study of the cell cycle has contributed to generate valuable ground-based knowledge of the biology of *Capsaspora*, we were still lacking methods to approach gene function at the molecular level of the DNA.

a) The value of the cell cycle in the development of model species

This thesis contributes to expanding the set of tools and knowledge available for one of the predilect species to study the origin of animal multicellularity, the filasterean *Capsaspora owczarzaki*. The study of *Capsaspora* at the genomic, transcriptional, proteomic and epigenomic level has helped to better understand genome evolution (and the evolution of genome regulation) under the framework of the evolutionary transition to animal multicellularity. Our novel data of periodic expression during the cell cycle adds to the seemingly ever-growing collection of resources surrounding *Capsaspora*, helping to elucidate and confirm previous assumptions regarding cell cycle regulation in Opisthokonta. Evolution and conservation of periodic gene expression might not have been a key factor to explain similarities and differences between animals and their unicellular relatives, due to the multiple additional layers of regulation controlling periodic activity during the cell cycle. On the contrary, expansion of cyclins and CDKs after the last common ancestor of animals and choanoflagellates has putative implications in the emergence of a more sophisticated control of the cell cycle responding to controlled, within-group proliferative signals. This scenario matches a situation where gene families expand, sub-functionalize, and acquire a specific niche of activation, a recurring motif in animal evolution (Paps and Holland, 2018; Richter et al., 2018). Overall, the independent

similarities in the periodic transcriptional program of *Capsaspora*, human cells, yeasts, and putatively other representatives from earlier-branching lineages such as *D. discoideum*, position *Capsaspora* as a new model system where to study cell cycle regulation in order to bridge gaps between our knowledge from yeasts and human cells.

Developing a model species involves getting to know the basis of its biology, serving as the foundations for further generation of knowledge. The cell cycle, being an intrinsic process to an organisms' biology, can inform about multiple aspects: from cell decision making, to duration and coordination of processes, to the mechanisms of DNA replication and mitosis. In addition, developing tools to study the cell cycle can have numerous side applications: culture synchronization allows the study of numerous cellular processes with a high temporal resolution and reproducibility. Similarly, accessing the cell's biology in a temporal manner allows to address questions regarding the duration of biological processes. Cells are known to react differently to certain treatments or extracellular cues according to their cell cycle phase, and show susceptibility to DNA uptake in the absence of nuclear membrane. This knowledge can help to design novel experimental protocols to address the species as a model organism.

b) Development of transfection in *Capsaspora*: technical breakthroughs and implications for functional studies

The biology of an organism can only be understood if its genome is both known (through genomics) and experimentally accessible (through genetic tools). In section 3.2, we show for the first time a robust protocol to transiently transfect Capsaspora adherent cells. Based on protocols previously used in other systems such as fibroblasts or *Dictyostelium* cells, we versioned a method of Calcium phosphate-based DNA precipitation to drive uptake of foreign DNA in *Capsaspora* adherent cells. The efficiency achieved is comparable to those from other transfection methods in other unicellular species (Booth et al., 2018; Caro et al., 2012; Gaudet et al., 2007; Chris J Janse et al., 2006; Chris J. Janse et al., 2006; Kawai et al., 2010; Schiestl and Gietz, 1989; Vinayak et al., 2015), and sufficient for further manipulations such as screenings and fluorescence/imaging approaches (cell imaging, subcellular tracking, live-cell immunofluorescence, phenotypic assays, or fluorescence-activated cell sorting (FACS)). As a result of our work, a thorough description of the protocol is available for reproducible testing, which has been successfully implemented in other laboratories achieving similar transfection efficiencies (Suga, Phillips, 2018 personal communications). Our protocol of DNA transfection is thus

the first step in a line of future genomic assays at the functional and molecular level; joining efforts with an antibiotic selection-based approach (Ros-Rocher, Parra-Acero et al., unpublished), our improved protocol of DNA transfection can contribute to building stable transfection and genome editing in Capsaspora.

c) Knowledge transfer across techniques in *Capsaspora*

Our protocols of DNA transfection and cell cycle synchronization have the potential to be used together, and progress in either sides might benefit each other. The ability to reproducibly transfect *Capsaspora* has contributed to our knowledge on very basic aspects of this organism, and might help to better understand its underlying biology. Tagging of plasma membrane revealed with unprecedented detail that *Capsaspora* cells attach to the substrate using their filopodia, with their cell bodies separated from the surface. This observation matches former descriptions of the movement of *Capsaspora* cells (Stibbs et al., 1979) and also matches our findings of cytoskeletal dynamics during the *Capsaspora* cell cycle, where cells upregulate actin remodelers during mitosis in parallel to downregulation of some components of the integrin adhesome. *Capsaspora* cells remain transfected for more than ten days, a time which exceeds the duration of other cellular processes such as the cell cycle (as characterized with our protocol) and thus undergoing multiple rounds of cell division. Therefore, transfection of *Capsaspora* with nuclear and cytoskeletal markers can, together with our protocol of culture synchronization, prove useful to *in vivo* monitor chromosome segregation and changes in the cytoskeletal dynamics.

Our knowledge of cell cycle regulation may allow the development of tools driven by cell cycle progression which, in turn, can help to further investigate and confirm former findings of the cell cycle; one potential application is engineering cell cycle-regulated constructs by clonings under the control of the promoters of periodic genes. Among other applications, transfection can help to confirm gene expression throughout the cell cycle, to track mitosis using markers of chromatin and nuclear membrane, to confirm cell cycle gene expression by promoters and transcription factors, to identify cells undergoing a specific cell cycle phase for further characterization, or to find some of the previously discussed cell cycle inhibitors by over-expressing putative candidates.

d) Knowledge transfer between *Capsaspora* and other species

Synergic progress not only occurs between techniques within a species but also across different species. Originally, Capsaspora transfection was inspired by previous protocols to transfect other amoeboid cells, such as D. discoideum and fibroblasts from mammals. The development of Capsaspora transfection has come hand in hand with the development of transfection in other unicellular Holozoa. This scenario mirrors the history of transgenesis development in unicellular algae (Jakobiak, 2004; Lerche and Hallmann, 2014, 2013, 2009), or the development of techniques across more traditional model organisms (reviewed in (Müller and Grossniklaus, 2010)). Our protocol of transfection in *Capsaspora* is contemporary to transfection and forward genetics in other lineages of unicellular relatives of animals, including choanoflagellates (Booth et al., 2018; Levin et al., 2014), ichthyosporeans (Suga and Ruiz-Trillo, 2013), and corallochytreans (Kozyczkowska et al., in preparation). Much like transfection in Capsaspora, such advances have contributed to our knowledge on fundamental aspects of the biology of these organisms, such as the localization and arrangement of proteins involved in the multicellular development of *S. rosetta*, or the synchronous nuclear division in *C. fragrantissima*.

In addition, our studies of culture synchronization and cell cycle progression can promote the initiative to investigate and compare the regulation of the cell cycle across multiple species of unicellular holozoans. Many of them exhibit unique life cycles of multinucleated cells (Glockling et al., 2013; Mendoza et al., 2002) or clonal colonies (Dayel et al., 2011) with similar behavior as some embryonic stages of animals. Studying their cell cycle regulation would unravel similarities between them, as well as have potential implications in the control of the cell cycle in the unicellular ancestor of animals. For example, internal duplications of CDK1 within choanoflagellates might be related to a more sophisticated regulation of the cell cycle related to their life cycles with uni- and multicellular stages, for which this lineage proves as an enticing point from where to continue this research.

This knowledge transfer across novel model organisms foretells the emergence of a new era of functional approaches in the field, being able for the first time to address long-standing questions about the role of so-thought “animal” genes in a unicellular context, including their contribution, if any, to the origin of multicellular organisms (Brunet and King, 2017; Richter and King, 2013; Rokas, 2008; Ruiz-Trillo et al., 2007).

4.3 Concluding remarks of this PhD thesis

Below lies a list of findings and concluding remarks from the results presented in this thesis, which illustrate the conclusions of section 5:

1. The regulation of the core cell cycle oscillator in *Capsaspora* shows similarities with animals, with a sequential order of expression of cyclins E, A, B, and CDK1 dynamically expressed throughout the cell cycle of *Capsaspora*.
2. The periodic transcriptional program of *Capsaspora* contains many genes involved in ancient machineries of DNA replication, chromatin assembly, and chromosome segregation, showing similarities to other model organisms studied such as yeasts or human cells.
3. Multiple transcription factors are dynamically expressed during the cell cycle of *Capsaspora*. A putative set of cell cycle transcriptional regulators in Capsaspora include a MAPK-Rb/E2F7-E2F/DP network of interaction controlling entry into cell cycle, and the transcriptional control of mitotic genes by an MYB-like transcription factor.
4. The absence of cell cycle regulators typical from G1 phase in other opisthokonts, such as CDK inhibitors Sic, p15, or p16, suggest that *Capsaspora* might control entry into the cell cycle differently from those organisms.
5. The periodic transcriptional program of *Capsaspora* contains a small but significant fraction of genes belonging to the same groups of orthology as human cells, yeasts, and one amoebozoan species. These similarities, together with a conserved set of CDK1, Cyclin A/B/E, E2F, and Rb orthologs, among other proteins absent from yeasts, position Capsaspora as a novel organism in which to compare cell cycle regulation within the group of Opisthokonta, known to have yielded complex multicellular forms independently.
6. Because cell cycle regulation can occur at multiple levels and evolves rapidly, this thesis illustrates the need to study more species outside the conventional catalog in order to study evolution and conservation of cell cycle regulation, and to develop genetic tools which may facilitate functional analysis.
7. Calcium-phosphate DNA precipitation can be used to transfect *Capsaspora* cells with an efficiency of ~1% 18 hours post transfection, a time sufficient for the functional assays during the cell cycle.
8. In the light of initial experiments, *Capsaspora* transfected cells can be synchronized and progress through the cell cycle normally. Our improved protocol of cell transfection could serve as the basis for functional assays of cell cycle genes in a unicellular relative of animals.

5. Conclusions

The main conclusions of the work presented in this thesis, related to the objectives listed in Section 2, are the following:

1. A hydroxyurea-based culture synchronization protocol allows the follow up of the cell cycle in *Capsaspora*.
2. Capsaspora owczarzaki shows a conventional cell cycle regulation consisting of a cyclin-CDK oscillator resembling animal regulation in terms of gene expression dynamics, which helps us to better understand how cyclin-CDK regulation evolved complexified in Opisthokonta.
3. Transcriptional regulation throughout the cell cycle of Capsaspora consists of a set of Rb/E2F/DP-MYB putatively regulating periodic gene expression, and a periodic transcriptional program containing conserved cell cycle regulators involved in DNA replication and mitosis checkpoints, including protein degradation regulated by the Anaphase-promoting complex.
4. The Capsaspora periodic transcriptional program has similarities to other unicellular and multicellular opisthokonts, which position this species as an enticing model system for comparative studies of the cell cycle.
5. We developed a calcium-phosphate transfection protocol in Capsaspora, which for the first time opens the door to functional assays and the future development of more genetic tools in a close unicellular relative of animals.

Bibliography

Adl SM, Bass D, Lane CE, Lukeš J, Schoch CL, Smirnov A, Agatha S, Berney C, Brown MW, Burki F, Cárdenas P, Čepička I, Chistyakova L, del Campo J, Dunthorn M, Edvardsen B, Eglit Y, Guillou L, Hampl V, Heiss AA, Hoppenrath M, James TY, Karnkowska A, Karpov S, Kim E, Kolisko M, Kudryavtsev A, Lahr DJG, Lara E, Le Gall L, Lynn DH, Mann DG, Massana R, Mitchell EAD, Morrow C, Park JS, Pawlowski JW, Powell MJ, Richter DJ, Rueckert S, Shadwick L, Shimano S, Spiegel FW, Torruella G, Youssef N, Zlatogursky V, Zhang Q. 2019. Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. *J Eukaryot Microbiol*. doi:10.1111/jeu.12691

Adl SM, Simpson AGB, Lane CE, Lukeš J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampl V, Heiss A, Hoppenrath M, Lara E, Gall L Le, Lynn DH, McManus H, Mitchell EAD, Mozley-Stanridge SE, Parfrey LW, Pawlowski J, Rueckert S, Shadwick L, Schoch CL, Smirnov A, Spiegel FW. 2012. The revised classification of eukaryotes. *J Eukaryot Microbiol* **59**:429–493. doi:10.1111/j.1550-7408.2012.00644.x

Ahmed S, Sheng H, Niu L, Henderson E. 1998. Tetrahymena mutants with short telomeres. *Genetics* **150**:643–650.

Alegado RA, Brown LW, Cao S, Dermenjian RK, Zuzow R, Fairclough SR, Clardy J, King N. 2012. A bacterial sulfonolipid triggers multicellular development in the closest living relatives of animals. *Elife* **2012**. doi:10.7554/eLife.00013

Araki S, Ito M, Soyano T, Nishihama R, Machida Y. 2004. Mitotic cyclins stimulate the activity of c-Myb-like factors for transactivation of G2/M phase-specific genes in tobacco. *J Biol Chem* **279**:32979–32988. doi:10.1074/jbc.M403171200

Archambault V, Glover DM. 2009. Polo-like kinases: Conservation and divergence in their functions and regulation. *Nat Rev Mol Cell Biol*. doi:10.1038/nrm2653

Aviner R, Geiger T, Elroy-Stein O. 2013. Novel proteomic approach (PUNCH-P) reveals cell cycle-specific fluctuations in mRNA translation. *Genes Dev* **27**:1834–1844. doi:10.1101/gad.219105.113

Bar-Joseph Z, Siegfried Z, Brandeis M, Brors B, Lu Y, Eils R, Dynlacht BD, Simon I.

2008. Genome-wide transcriptional analysis of the human cell cycle identifies genes differentially regulated in normal and cancer cells. *Proc Natl Acad Sci*. doi:10.1073/pnas.0704723105

Bar-On YM, Phillips R, Milo R. 2018. The biomass distribution on Earth. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.1711842115

Barbier M, Leighfield TA, Soyer-Gobillard MO, Van Dolah FM. 2003. Permanent expression of a cyclin B homologue in the cell cycle of the dinoflagellate Karenia brevis. *J Eukaryot Microbiol* **50**:123–131. doi:10.1111/j.1550-7408.2003.tb00246.x

Barron M, Li J. 2016. Identifying and removing the cell-cycle effect from single-cell RNA-Sequencing data. *Sci Rep* **6**. doi:10.1038/srep33892

Barry C, Li X, Kendziorski C, Thomson JA, Choi J, Chu L-F, Stewart RM, Jiang P, Leng N, Li Y. 2015. Oscope identifies oscillatory genes in unsynchronized single-cell RNA-seq experiments. *Nat Methods* **12**:947–950. doi:10.1038/nmeth.3549

Bateman A, Martin MJ, O'Donovan C, Magrane M, Alpi E, Antunes R, Bely B, Bingley M, Bonilla C, Britto R, Bursteinas B, Bye-AJee H, Cowley A, Da Silva A, De Giorgi M, Dogan T, Fazzini F, Castro LG, Figueira L, Garmiri P, Georghiou G, Gonzalez D, Hatton-Ellis E, Li W, Liu W, Lopez R, Luo J, Lussi Y, MacDougall A, Nightingale A, Palka B, Pichler K, Poggioli D, Pundir S, Pureza L, Qi G, Rosanoff S, Saidi R, Sawford T, Shypitsyna A, Speretta E, Turner E, Tyagi N, Volynkin V, Wardell T, Warner K, Watkins X, Zaru R, Zellner H, Xenarios I, Bougueleret L, Bridge A, Poux S, Redaschi N, Aimo L, ArgoudPuy G, Auchincloss A, Axelsen K, Bansal P, Baratin D, Blatter MC, Boeckmann B, Bolleman J, Boutet E, Breuza L, Casal-Casas C, De Castro E, Coudert E, Cuche B, Doche M, Dornevil D, Duvaud S, Estreicher A, Famiglietti L, Feuermann M, Gasteiger E, Gehant S, Gerritsen V, Gos A, Gruaz-Gumowski N, Hinz U, Hulo C, Jungo F, Keller G, Lara V, Lemercier P, Lieberherr D, Lombardot T, Martin X, Masson P, Morgat A, Neto T, Nouspikel N, Paesano S, Pedruzzi I, Pilbout S, Pozzato M, Pruess M, Rivoire C, Roechert B, Schneider M, Sigrist C, Sonesson K, Staehli S, Stutz A, Sundaram S, Tognolli M, Verbregue L, Veuthey AL, Wu CH, Arighi CN, Arminski L, Chen C, Chen Y, Garavelli JS, Huang H, Laiho K, McGarvey P, Natale DA, Ross K, Vinayaka CR, Wang Q, Wang Y, Yeh LS, Zhang J. 2017. UniProt: The universal protein knowledgebase. *Nucleic Acids Res*. doi:10.1093/nar/gkw1099

Beach D, Serrano M, Hannon GJ. 1993. inhibition of cyclin D / CDK4 **366**:704–707.

Ben-David U, Ha G, Tseng Y-Y, Greenwald NF, Oh C, Shih J, McFarland JM, Wong B, Boehm JS, Beroukhim R, Golub TR. 2017. Patient-derived xenografts undergo mouse-specific tumor evolution. *Nat Genet* **49**:1567.

Ben-David U, Siranosian B, Ha G, Tang H, Oren Y, Hinohara K, Strathdee CA, Dempster J, Lyons NJ, Burns R, Nag A, Kugener G, Cimini B, Tsvetkov P, Maruvka YE, O'Rourke R, Garrity A, Tubelli AA, Bandopadhayay P, Tsherniak A, Vazquez F, Wong B, Birger C, Ghandi M, Thorner AR, Bittker JA, Meyerson M, Getz G, Beroukhim R, Golub TR. 2018. Genetic and transcriptional evolution alters cancer cell line drug response. *Nature* **560**:325–330. doi:10.1038/s41586-018-0409-3

Bendris N, Lemmers B, Blanchard JM. 2015. Cell cycle, cytoskeleton dynamics and beyond: the many functions of cyclins and CDK inhibitors. *Cell Cycle* **14**:1786–1798. doi:10.1080/15384101.2014.998085

Bhaud Y, Guillebault D, Lennon JF, Defacque H, Soyer-Gobillard MO, Moreau H. 2000. Morphology and behaviour of dinoflagellate chromosomes during the cell cycle and mitosis. *J Cell Sci* **113**:1231–1239.

Bi E, Park HO. 2012. Cell polarization and cytokinesis in budding yeast. *Genetics* **191**:347–387. doi:10.1534/genetics.111.132886

Bloom J, Cross FR. 2007. Multiple levels of cyclin specificity in cell-cycle control. *Nat Rev Mol Cell Biol*. doi:10.1038/nrm2105

Boeck ME, Huynh C, Gevirtzman L, Thompson OA, Wang G, Kasper DM, Reinke V, Hillier LW, Waterston RH. 2016. The time-resolved transcriptome of C. Elegans. *Genome Res*. doi:10.1101/gr.202663.115

Bonke M, Turunen M, Sokolova M, Vähärautio A, Kivioja T, Taipale M, Björklund M, Taipale J. 2013. Transcriptional Networks Controlling the Cell Cycle. *G3 - Genes|Genomes|Genetics* **3**:75–90. doi:10.1534/g3.112.004283

Bonner JT. 2009. First Signals, First Signals: The Evolution of Multicellular Development. Princeton: Princeton University Press. doi:10.1515/9781400830589

Booher R, Beach D. 1988. Involvement of cdc13+ in mitotic control in

Schizosaccharomyces pombe: possible interaction of the gene product with microtubules. *EMBO J*. doi:10.1002/j.1460-2075.1988.tb03075.x

Booher RN, Alfa CE, Hyams JS, Beach DH. 1989. The fission yeast cdc2/cdc13/suc1 protein kinase: Regulation of catalytic activity and nuclear localization. *Cell*. doi:10.1016/0092-8674(89)90429-7

Booth DS, Szmidt-Middleton H, King N. 2018. Transfection of choanoflagellates illuminates their cell biology and the ancestry of animal septins. *Mol Biol Cell* **29**:3026–3038. doi:10.1091/mbc.E18-08-0514

Boraas ME, Seale DB, Boxhorn JE. 1998. Phagotrophy by flagellate selects for colonial prey: A possible origin of multicellularity. *Evol Ecol* **12**:153–164. doi:10.1023/A:1006527528063

Borel F, Lacroix FB, Margolis RL. 2002. Prolonged arrest of mammalian cells at the G1/S boundary results in permanent S phase stasis. *J Cell Sci* **115**:2829–38.

Bosher JM, Labouesse M. 2000. RNA interference: genetic wand and genetic watchdog. *Nat Cell Biol* **2**:E31–E36. doi:10.1038/35000102

Botz JR, Zerfass-Thome K, Spitkovsky D, Delius H, Vogt B, Eilers M, Hatzigeorgiou A, Pidder A, Rr J-D, Tumorvirologie FA, Bioinformatik F. 1996. Cell Cycle Regulation of the Murine Cyclin E Gene Depends on an E2F Binding Site in the Promoter Cyclin E controls progression through the G 1 phase of the cell cycle in mammalian fibroblasts and potentially in many other cell types. Cyclin E is a rate-li. *Mol Cell Biol* **16**:3401–3409.

Breyne P, Dreesen R, Vandepoele K, De Veylder L, Van Breusegem F, Callewaert L, Rombauts S, Raes J, Cannoot B, Engler G, Inze D, Zabeau M. 2002. Transcriptome analysis during cell division in plants. *Proc Natl Acad Sci* **99**:14825–14830. doi:10.1073/pnas.222561199

Brown MW, Silberman JD. 2013. The Non-dictyostelid Sorocarpic AmoebaeDictyostelids. Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 219–242. doi:10.1007/978-3-642-38487-5_12

Brown NR, Korolchuk S, Martin MP, Stanley WA, Moukhametzianov R, Noble MEM, Endicott JA. 2015. CDK1 structures reveal conserved and unique features of the

essential cell cycle CDK. *Nat Commun* **6**:1–12. doi:10.1038/ncomms7769

Brown NR, Noble M, Endicott JA, Garman EF, Wakatsuki S, Mitchell E, Rasmussen B, Hunt T, Johnson LN. 1995. The crystal structure of cyclin A. *Structure* **3**:1235–1247. doi:10.1016/S0969-2126(01)00259-3

Brunelle SA, Van Dolah FM. 2011. Post-transcriptional regulation of S-Phase genes in the dinoflagellate, karenia brevis. *J Eukaryot Microbiol* **58**:373–382. doi:10.1111/j.1550-7408.2011.00560.x

Brunet T, King N. 2017. The Origin of Animal Multicellularity and Cell Differentiation. *Dev Cell* **43**:124–140. doi:10.1016/j.devcel.2017.09.016

Buss LW. 1988. The evolution of individuality. *Trends Ecol Evol* **3**:214. doi:10.1016/0169-5347(88)90010-9

Cao L, Chen F, Yang X, Xu W, Xie J, Yu L. 2014. Phylogenetic analysis of CDK and cyclin proteins in premetazoan lineages. *BMC Evol Biol*. doi:10.1186/1471-2148-14-10

Caretti G, Salsi V, Vecchi C, Imbriano C, Mantovani R. 2003. Dynamic recruitment of NF-Y and histone acetyltransferases on cell-cycle promoters. *J Biol Chem*. doi:10.1074/jbc.M304606200

Caro F, Miller MG, DeRisi JL. 2012. Plate-based transfection and culturing technique for genetic manipulation of Plasmodium falciparum. *Malar J* **11**:22. doi:10.1186/1475-2875-11-22

CAVALIER-SMITH T. 1987. The Origin of Eukaryote and Archaebacterial Cells. *Ann N Y Acad Sci*. doi:10.1111/j.1749-6632.1987.tb40596.x

Chen EH, Grote E, Mohler W, Vignery A. 2007. Cell-cell fusion. *FEBS Lett* **581**:2181–2193. doi:10.1016/j.febslet.2007.03.033

Cho RJ, Huang M, Campbell MJ, Dong H, Steinmetz L, Sapinoso L, Hampton G, Elledge SJ, Davis RW, Lockhart DJ. 2001. Transcriptional regulation and function during the human cell cycle. *Nat Genet*. doi:10.1038/83751

Clark HJ. 1866. Conclusive proofs of the animality of the ciliate sponges, and of their

affinities with the Infusoria flagellata. *Am J Sci* **s2-42**:320–324. doi:10.2475/ajs.s2-42.126.320

Clarke E. 2014. Origins of evolutionary transitions. *J Biosci*. doi:10.1007/s12038-013-9375-y

Cohen-Fix O, Peters JM, Kirschner MW, Koshland D. 1996. Anaphase initiation in saccharomyces cerevisiae is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p. *Genes Dev*. doi:10.1101/gad.10.24.3081

Cook JG, Park CH, Burke TW, Leone G, DeGregori J, Engel A, Nevins JR. 2002. Analysis of Cdc6 function in the assembly of mammalian prereplication complexes. *Proc Natl Acad Sci U S A* **99**:1347–1352. doi:10.1073/pnas.032677499

Cooper K. 2006. Rb, whi it's not just for metazoans anymore. *Oncogene* **25**:5228–5232. doi:10.1038/sj.onc.1209630

Costa RH. 2005. FoxM1 dances with mitosis. *Nat Cell Biol* **7**:108–110. doi:10.1038/ncb0205-108

Côte P, Hogues H, Whiteway M. 2009. Transcriptional Analysis of the Candida albicans Cell Cycle. *Mol Biol Cell* **20**:3363–3373. doi:10.1091/mbc.E09

Coudreuse D, Nurse P. 2010. Driving the cell cycle with a minimal CDK control network. *Nature*. doi:10.1038/nature09543

Courjol F, Gissot M. 2018. A coiled-coil protein is required for coordination of karyokinesis and cytokinesis in Toxoplasma gondii. *Cell Microbiol* **20**. doi:10.1111/cmi.12832

Coverley D, Laman H, Laskey RA. 2002. Distinct roles for cyclins E and A during DNA replication complex assembly and activation. *Nat Cell Biol*. doi:10.1038/ncb813

Cross FR, Buchler NE, Skotheim JM. 2011. Evolution of networks and sequences in eukaryotic cell cycle control. *Philos Trans R Soc B Biol Sci* **366**:3532–3544. doi:10.1098/rstb.2011.0078

Cross FR, Tinkelenberg AH. 1991. A potential positive feedback loop controlling CLN1 and CLN2 gene expression at the start of the yeast cell cycle. *Cell* **65**:875–883.

doi:10.1016/0092-8674(91)90394-E

Cross FR, Yuste-Rojas M, Gray S, Jacobson MD. 1999. Specialization and targeting of B-type cyclins. *Mol Cell* **4**:11–19. doi:10.1016/S1097-2765(00)80183-5

Crowley TE, Nellen W, Gomer RH, Firtel RA. 1985. Phenocopy of discoidin I-minus mutants by antisense transformation in Dictyostelium. *Cell* **43**:633–641. doi:10.1016/0092-8674(85)90235-1

Dayel MJ, Alegado RA, Fairclough SR, Levin TC, Nichols SA, McDonald K, King N. 2011. Cell differentiation and morphogenesis in the colony-forming choanoflagellate Salpingoeca rosetta. *Dev Biol* **357**:73–82. doi:10.1016/j.ydbio.2011.06.003

De Bondt HL, Rosenblatt J, Jancarik J, Jones HD, Morgant DO, Kim SH. 1993. Crystal structure of cyclin-dependent kinase 2. *Nature* **363**:595–602. doi:10.1038/363595a0

De Melo Godoy PD, Nogueira-Junior LA, Paes LS, Cornejo A, Martins RM, Silber AM, Schenkman S, Elias MC. 2009. Trypanosome prereplication machinery contains a single functional Orc1/Cdc6 protein, which is typical of archaea. *Eukaryot Cell* **8**:1592–1603. doi:10.1128/EC.00161-09

de Mendoza A, Sebé-Pedrós A. 2019. Origin and evolution of eukaryotic transcription factors. *Curr Opin Genet Dev* **58**–**59**:25–32. doi:10.1016/j.gde.2019.07.010

de Mendoza A, Sebe-Pedros A, Sestak MS, Matejcic M, Torruella G, Domazet-Loso T, Ruiz-Trillo I. 2013. Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proc Natl Acad Sci* **110**:E4858–E4866. doi:10.1073/pnas.1311818110

de Mendoza A, Suga H, Permanyer J, Irimia M, Ruiz-Trillo I. 2015. Complex transcriptional regulation and independent evolution of fungal-like traits in a relative of animals. *Elife*. doi:10.7554/eLife.08904.001

De Monte S, Rainey PB. 2014. Nascent multicellular life and the emergence of individuality. *J Biosci* **39**:237–248. doi:10.1007/s12038-014-9420-5

del Campo J, Mallo D, Massana R, de Vargas C, Richards TA, Ruiz-Trillo I. 2015. Diversity and distribution of unicellular opisthokonts along the European coast

analysed using high-throughput sequencing. *Environ Microbiol* **17**:3195–3207. doi:10.1111/1462-2920.12759

del Campo J, Ruiz-Trillo I. 2013. Environmental Survey Meta-analysis Reveals Hidden Diversity among Unicellular Opisthokonts. *Mol Biol Evol* **30**:802–805. doi:10.1093/molbev/mst006

Del Campo J, Sieracki ME, Molestina R, Keeling P, Massana R, Ruiz-Trillo I. 2014. The others: Our biased perspective of eukaryotic genomes. *Trends Ecol Evol*. doi:10.1016/j.tree.2014.03.006

Den Haese GJ, Walworth N, Carr AM, Gould KL. 1995. The Wee1 protein kinase regulates T14 phosphorylation of fission yeast Cdc2. *Mol Biol Cell* **6**:371–385. doi:10.1091/mbc.6.4.371

Desvoyes B, de Mendoza A, Ruiz-Trillo I, Gutierrez C. 2014. Novel roles of plant RETINOBLASTOMA-RELATED (RBR) protein in cell proliferation and asymmetric cell division. *J Exp Bot* **65**:2657–2666. doi:10.1093/jxb/ert411

Dewachter L, Verstraeten N, Fauvart M, Michiels J. 2018. An integrative view of cell cycle control in Escherichia coli. *FEMS Microbiol Rev*. doi:10.1093/femsre/fuy005

Doenecke D. 2016. Chromatin Dynamics During the Cell CycleChromatin Regulation and Dynamics. doi:10.1016/B978-0-12-803395-1.00005-8

Dominguez D, Tsai YH, Gomez N, Jha DK, Davis I, Wang Z. 2016. A high-resolution transcriptome map of cell cycle reveals novel connections between periodic genes and cancer. *Cell Res* **26**:946–962. doi:10.1038/cr.2016.84

Donachie W. 1993. The Cell Cycle of Escherichia coll. *Annu Rev Microbiol* **47**:199–230. doi:10.1146/annurev.micro.47.1.199

Donczew M, Mackiewicz P, Wróbel A, Flärdh K, Zakrzewska-Czerwińska J, Jakimowicz D. 2016. ParA and ParB coordinate chromosome segregation with cell elongation and division during Streptomyces sporulation. *Open Biol* **6**. doi:10.1098/rsob.150263

Dorin-Semblat D, Carvalho TG, Nivez M-P, Halbert J, Poullet P, Semblat J-P, Goldring D, Chakrabarti D, Mehra P, Dhar S, Paing MM, Goldberg DE, McMillan PJ, Tilley L,

Doerig C. 2013. An atypical cyclin-dependent kinase controls Plasmodium falciparum proliferation rate. *Kinome* **1**. doi:10.2478/kinome-2013-0001

Dove WF, Dee J, Hatani S, Haugli FB, Wohlfarth-Bottermann K-E. 1986. The Molecular Biology of Physarum polycephalum, The Molecular Biology of Physarum polycephalum. Springer New York. doi:10.1007/978-1-4613-2203-0

Draetta G, Luca F, Westendorf J, Brizuela L, Ruderman J, Beach D. 1989. cdc2 protein kinase is complexed with both cyclin A and B: Evidence for proteolytic inactivation of MPF. *Cell* **56**:829–838. doi:10.1016/0092-8674(89)90687-9

Dudin O, Ondracka A, Grau-Bove X, Haraldsen AAB, Toyoda A, Suga H, Brate J, Ruiz-Trillo I. 2019. A Unicellular Relative of Animals Generates an Epithelium-Like Cell Layer by Actomyosin-dependent Cellularization. *SSRN Electron J*. doi:10.2139/ssrn.3362312

Duffy KT, McAleer MF, Davidson WR, Kari L, Kari C, Liu CG, Farber SA, Cheng KC, Mest JR, Wickstrom E, Dicker AP, Rodeck U. 2005. Coordinate control of cell cycle regulatory genes in zebrafish development tested by cyclin D1 knockdown with morpholino phosphorodiamidates and hydroxyprolyl-phosphono peptide nucleic acids. *Nucleic Acids Res*. doi:10.1093/nar/gki799

Dunphy WG, Newport JW. 1988. Mitosis-inducing factors are present in a latent form during interphase in the Xenopus embryo. *J Cell Biol* **106**:2047–2056. doi:10.1083/jcb.106.6.2047

Eme L, Spang A, Lombard J, Stairs CW, Ettema TJG. 2017. Archaea and the origin of eukaryotes. *Nat Rev Microbiol* **15**:711–723. doi:10.1038/nrmicro.2017.133

Eme L, Trilles A, Moreira D, Brochier-Armanet C. 2011. The phylogenomic analysis of the anaphase promoting complex and its targets points to complex and modern-like control of the cell cycle in the last common ancestor of eukaryotes. *BMC Evol Biol*. doi:10.1186/1471-2148-11-265

Epstein CB, Cross FR. 1992. CLB5: A novel B cyclin from budding yeast with a role in S phase. *Genes Dev* **6**:1695–1706. doi:10.1101/gad.6.9.1695

Evans T, Rosenthal ET, Youngblom J, Distel D, Hunt T. 1983. Cyclin: A protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division.

Cell **33**:389–396. doi:10.1016/0092-8674(83)90420-8

Fairclough SR, Chen Z, Kramer E, Zeng Q, Young S, Robertson HM, Begovic E, Richter DJ, Russ C, Westbrook MJ, Manning G, Lang BF, Haas B, Nusbaum C, King N. 2013. Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate Salpingoeca rosetta. *Genome Biol* **14**:1–15. doi:10.1186/gb-2013-14-2-r15

Fang G, Hongtao Y, Kirschner MW. 1998. Direct binding of CDC20 protein family members activates the anaphase-promoting complex in mitosis and G1. *Mol Cell* **2**:163–171. doi:10.1016/S1097-2765(00)80126-4

Farrell JA, O'Farrell PH. 2014. From Egg to Gastrula: How the Cell Cycle Is Remodeled During the Drosophila Mid-Blastula Transition . *Annu Rev Genet* **48**:269–294. doi:10.1146/annurev-genet-111212-133531

Fay P, Walsby AE. 1966. Metabolic activities of isolated heterocysts of the blue-green alga Anabaena cylindrica [36]. *Nature*. doi:10.1038/209094a0

Ferguson DJP, Sahoo N, Pinches RA, Bumstead JM, Tomley FM, Gubbels MJ. 2008. MORN1 has a conserved role in asexual and sexual development across the Apicomplexa. *Eukaryot Cell* **7**:698–711. doi:10.1128/EC.00021-08

Ferrer-Bonet M, Ruiz-Trillo I. 2017. Capsaspora owczarzaki. *Curr Biol*. doi:10.1016/j.cub.2017.05.074

Fey P, Kowal AS, Gaudet P, Pilcher KE, Chisholm RL. 2007. Protocols for growth and development of Dictyostelium discoideum. *Nat Protoc* **2**:1307–1316. doi:10.1038/nprot.2007.178

Fischer M, DeCaprio JA. 2015. Does Arabidopsis thaliana DREAM of cell cycle control? . *EMBO J* **34**:1987–1989. doi:10.15252/embj.201592196

Fischer M, Müller GA. 2017. Cell cycle transcription control: DREAM/MuvB and RB-E2F complexes. *Crit Rev Biochem Mol Biol* **52**:638–662. doi:10.1080/10409238.2017.1360836

Fisher D, Krasinska L, Coudreuse D, Novák B. 2012. Phosphorylation network dynamics in the control of cell cycle transitions. *J Cell Sci* **125**:4703–4711.

doi:10.1242/jcs.106351

Francia ME, Striepen B. 2014. Cell division in apicomplexan parasites. *Nat Rev Microbiol*. doi:10.1038/nrmicro3184

Fukui Y, Inoue S. 1991. Cell division in Dictyostelium with special emphasis on actomyosin organization in cytokinesis. *Cell Motil Cytoskeleton* **18**:41–54. doi:10.1002/cm.970180105

Gallone B, Steensels J, Prahl T, Soriaga L, Saels V, Herrera-Malaver B, Merlevede A, Roncoroni M, Voordeckers K, Miraglia L, Teiling C, Steffy B, Taylor M, Schwartz A, Richardson T, White C, Baele G, Maere S, Verstrepen KJ. 2016. Domestication and Divergence of Saccharomyces cerevisiae Beer Yeasts. *Cell* **166**:1397-1410.e16. doi:10.1016/j.cell.2016.08.020

Ganter M, Goldberg JM, Dvorin JD, Paulo JA, King JG, Tripathi AK, Paul AS, Yang J, Coppens I, Jiang RHY, Elsworth B, Baker DA, Dinglasan RR, Gygi SP, Duraisingh MT. 2017. Erratum: Plasmodium falciparum CRK4 directs continuous rounds of DNA replication during schizogony. *Nat Microbiol* **2**:17038. doi:10.1038/nmicrobiol.2017.38

Garcia K, Duncan T, Su TT. 2007. Analysis of the cell division cycle in Drosophila. *Methods* **41**:198–205. doi:10.1016/j.ymeth.2006.08.013

Gaudet P, Pilcher KE, Fey P, Chisholm RL. 2007. Transformation of Dictyostelium discoideum with plasmid DNA. *Nat Protoc* **2**:1317–1324. doi:10.1038/nprot.2007.179

Gefeng Z, Spellman PT, Volpe T, Brown PO, Botstein D, Davis TN, Futcher B. 2000. Two yeast forkhead genes regulate the cell cycle and pseudohyphal growth. *Nature* **406**:90–94. doi:10.1038/35017581

Gerber T, Murawala P, Knapp D, Masselink W, Schuez M, Hermann S, Gac-Santel M, Nowoshilow S, Kageyama J, Khattak S, Currie JD, Camp JG, Tanaka EM, Treutlein B. 2018. Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. *Science* **362**. doi:10.1126/science.aaq0681

Gibson TJ, Thompson JD, Blocker A, Kouzarides T. 1994. Evidence for a protein domain superfamily shared by the cyclins, TFIIB and RB/p107. *Nucleic Acids Res* **22**:946–

952. doi:10.1093/nar/22.6.946

Giotti B, Joshi A, Freeman TC. 2017. Meta-analysis reveals conserved cell cycle transcriptional network across multiple human cell types. *BMC Genomics*. doi:10.1186/s12864-016-3435-2

Glockling SL, Marshall WL, Gleason FH. 2013. Phylogenetic interpretations and ecological potentials of the Mesomycetozoea (Ichthyosporea). *Fungal Ecol*. doi:10.1016/j.funeco.2013.03.005

Golan A, Yudkovsky Y, Hershko A. 2002. The cyclin-ubiquitin ligase activity of cyclosome/APC is jointly activated by protein kinases Cdk1-cyclin B and Plk. *J Biol Chem* **277**:15552–15557. doi:10.1074/jbc.M111476200

Goldstein B, King N. 2016. The Future of Cell Biology: Emerging Model Organisms. *Trends Cell Biol* **26**:818–824. doi:10.1016/j.tcb.2016.08.005

Gomer RH, Firtel RA. 1987. Cell-autonomous determination of cell-type choice in Dictyostelium development by cell-cycle phase. *Science* **237**:758–762. doi:10.1126/science.3039657

Goodrich DW, Wang NP, Qian YW, Lee EYHP, Lee WH. 1991. The retinoblastoma gene product regulates progression through the G1 phase of the cell cycle. *Cell*. doi:10.1016/0092-8674(91)90181-W

Goss R. 1969. Principles of Regeneration, 1st ed. USA: Elsevier Books.

Gourguechon S, Wang CC. 2009. CRK9 contributes to regulation of mitosis and cytokinesis in the procyclic form of Trypanosoma brucei. *BMC Cell Biol* **10**:68. doi:10.1186/1471-2121-10-68

Graham FL, van der Eb AJ. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**:456–467. doi:10.1016/0042-6822(73)90341-3

Grant GD, Brooks L, Zhang X, Mahoney JM, Martyanov V, Wood TA, Sherlock G, Cheng C, Whitfield ML. 2013. Identification of cell cycle-regulated genes periodically expressed in U2OS cells and their regulation by FOXM1 and E2F transcription factors. *Mol Biol Cell* **24**:3634–3650. doi:10.1091/mbc.E13-05-0264

Grau-Bové X, Torruella G, Donachie S, Suga H, Leonard G, Richards TA, Ruiz-Trillo I. 2017. Dynamics of genomic innovation in the unicellular ancestry of animals. *Elife*. doi:10.7554/eLife.26036

Graumann PL. 2014. Chromosome architecture and segregation in prokaryotic cells. *J Mol Microbiol Biotechnol* **24**:291–300. doi:10.1159/000369100

Gray MW, Lukeš J, Archibald JM, Keeling PJ, Doolittle WF. 2010. Irremediable complexity? *Science* **330**:920–921. doi:10.1126/science.1198594

Griesemer J. 2001. The Units of Evolutionary Transition. *Selection*. doi:10.1556/select.1.2000.1-3.7

Grosberg RK, Strathmann RR. 2007. The Evolution of Multicellularity: A Minor Major Transition? *Annu Rev Ecol Evol Syst* **38**:621–654. doi:10.1146/annurev.ecolsys.36.102403.114735

Guertin DA, Trautmann S, McCollum D. 2002. Cytokinesis in Eukaryotes. *Microbiol Mol Biol Rev* **66**:155–178. doi:10.1128/mmbr.66.2.155-178.2002

Gunbin K V., Suslov V V., Turnaev II, Afonnikov DA, Kolchanov NA. 2011. Molecular evolution of cyclin proteins in animals and fungi. *BMC Evol Biol* **11**. doi:10.1186/1471-2148-11-224

Guo K, Walsh K. 1997. Inhibition of myogenesis by multiple cyclin-Cdk complexes: Coordinate regulation of myogenesis and cell cycle activity at the level of E2F. *J Biol Chem* **272**:791–797. doi:10.1074/jbc.272.2.791

Gutierrez C. 2009. The Arabidopsis Cell Division Cycle. *Arab B*. doi:10.1199/tab.0120

Guttes E, Guttes S. 1964. Chapter 3 Mitotic Synchrony in the Plasmodia of Physarum polycephalum and Mitotic Synchronization by Coalescence of Microplasmodia. *Methods Cell Biol* **1**:43–54. doi:10.1016/S0091-679X(08)62085-3

Guttes E, Guttes S, Rusch HP. 1961. Morphological observations on growth and differentiation of Physarum polycephalum grown in pure culture. *Dev Biol*. doi:10.1016/0012-1606(61)90034-3

Haber JE. 1998. MATING-TYPE GENE SWITCHING IN SACCHAROMYCES

CEREVISIAE. *Annu Rev Genet* **32**:561–599. doi:10.1146/annurev.genet.32.1.561

Hackett JD, Anderson DM, Erdner DL, Bhattacharya D. 2004. Dinoflagellates: A remarkable evolutionary experiment. *Am J Bot*. doi:10.3732/ajb.91.10.1523

Hagan IM, Grallert A, Simanis V. 2016. Analysis of the Schizosaccharomyces pombe cell cycle. *Cold Spring Harb Protoc*. doi:10.1101/pdb.top082800

Hammarton TC, Clark J, Douglas F, Boshart M, Mottram JC. 2003. Stage-specific differences in cell cycle control in Trypanosoma brucei revealed by RNA interference of a mitotic cyclin. *J Biol Chem* **278**:22877–22886. doi:10.1074/jbc.M300813200

Hammarton TC, Engstler M, Mottram JC. 2004. The Trypanosoma brucei cyclin, CYC2, is required for cell cycle progression through G1 phase and for maintenance of procyclic form cell morphology. *J Biol Chem* **279**:24757–24764. doi:10.1074/jbc.M401276200

Hanschen ER, Marriage TN, Ferris PJ, Hamaji T, Toyoda A, Fujiyama A, Neme R, Noguchi H, Minakuchi Y, Suzuki M, Kawai-Toyooka H, Smith DR, Sparks H, Anderson J, Bakarić R, Luria V, Karger A, Kirschner MW, Durand PM, Michod RE, Nozaki H, Olson BJSCSC. 2016. The Gonium pectorale genome demonstrates co-option of cell cycle regulation during the evolution of multicellularity. *Nat Commun* **7**:11370. doi:10.1038/ncomms11370

Hansen D V., Loktev A V., Ban KH, Jackson PK. 2004. Plk1 regulates activation of the anaphase promoting complex by phosphorylating and triggering SCFβTrCP_ dependent destruction of the APC inhibitor Emi1. *Mol Biol Cell*. doi:10.1091/mbc.E04-07-0598

Hartwell LH, Culotti J, Pringle JR, Reid BJ. 1974. Genetic Control of the Cell Division Cycle in Yeast: A model to account for the order of cell cycle events is deduced from the phenotypes of yeast mutants. *Science* **183**:46–51. doi:10.1126/science.183.4120.46

Heber-Katz E, Zhang Y, Bedelbaeva K, Song F, Chen X, Stocum DL. 2012. Cell Cycle Regulation and Regeneration. pp. 253–276. doi:10.1007/82_2012_294

Hehenberger E, Tikhonenkov D V., Kolisko M, del Campo J, Esaulov AS, Mylnikov AP,

Keeling PJ. 2017. Novel Predators Reshape Holozoan Phylogeny and Reveal the Presence of a Two-Component Signaling System in the Ancestor of Animals. *Curr Biol*. doi:10.1016/j.cub.2017.06.006

Helin K, Wu CL, Fattaey AR, Lees JA, Dynlacht BD, Ngwu C, Harlow E. 1993. Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperative trans-activation. *Genes Dev* **7**:1850–1861. doi:10.1101/gad.7.10.1850

Herron MD. 2016. Origins of multicellular complexity: Volvox and the volvocine algae. *Mol Ecol* **25**:1213–1223. doi:10.1111/mec.13551

Herron MD, Borin JM, Boswell JC, Walker J, Chen ICK, Knox CA, Boyd M, Rosenzweig F, Ratcliff WC. 2019. De novo origins of multicellularity in response to predation. *Sci Rep* **9**:2328. doi:10.1038/s41598-019-39558-8

Herron MD, Ratcliff WC, Boswell J, Rosenzweig F. 2018. Genetics of a de novo origin of undifferentiated multicellularity. *R Soc Open Sci* **5**. doi:10.1098/rsos.180912

Hertel LA, Bayne CJ, Loker ES. 2002. The symbiont Capsaspora owczarzaki, nov. gen. nov. sp., isolated from three strains of the pulmonate snail Biomphalaria glabrata is related to members of the Mesomycetozoea. *Int J Parasitol* **32**:1183–1191. doi:10.1016/S0020-7519(02)00066-8

Hong CI, Zamborszky J, Baek M, Labiscsak L, Ju K, Lee H, Larrondo LF, Goity A, Chong HS, Belden WJ, Csikasz-Nagy A. 2014. Circadian rhythms synchronize mitosis in Neurospora crassa. *Proc Natl Acad Sci* **111**:1397–1402. doi:10.1073/pnas.1319399111

Horowitz NH, Bonner D, Mitchell HK, Tatum EL, Beadle GW. 1945. Genic Control of Biochemical Reactions in Neurospora. *Am Nat* **79**:304–317. doi:10.1086/281267

Hunt T. 1989. Maturation promoting factor, cyclin and the control of M-phase. *Curr Opin Cell Biol* **1**:268–274. doi:10.1016/0955-0674(89)90099-9

Ingham PW. 1997. Zebrafish genetics and its implications for understanding vertebrate development. *Hum Mol Genet*. doi:10.1093/hmg/6.10.1755

Ishida S, Huang E, Zuzan H, Spang R, Leone G, West M, Nevins JR. 2001. Role for E2F in control of both DNA replication and mitotic functions as revealed from DNA

microarray analysis. *Mol Cell Biol* **21**:4684–4699. doi:10.1128/MCB.21.14.4684

Ito M, Araki S, Matsunaga S, Itoh T, Nishihama R, Machida Y, Doonan JH, Watanabe A. 2001. G2/M-phase-specific transcription during the plant cell cycle is mediated by c-Myb-like transcription factors. *Plant Cell* **13**:1891–1905. doi:10.1105/tpc.13.8.1891

Ivanova T, Maier M, Missarova A, Ziegler-Birling C, Carey LB, Mendoza M. 2018. Budding yeast complete DNA replication after chromosome segregation begins. *bioRxiv*. doi:10.1101/407957

Iyer VR, Eisen MB, Ross DT, Schuler G, Moore T, Lee JCF, Trent JM, Staudt LM, Hudson J, Boguski MS, Lashkari D, Shalon D, Botstein D, Brown PO. 1999. The transcriptional program in the response of human fibroblasts to serum. *Science* **283**:83–87. doi:10.1126/science.283.5398.83

Iyer VR, Horak CE, Scafe CS, Botstein D, Snyder M, Brown PO. 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**:533–538. doi:10.1038/35054095

Jakobiak T. 2004. The bacterial paromomycin resistance gene, aphH, as a dominant selectable marker in Volvox carteri. *Protist* **155**:381–393. doi:10.1078/1434461042650343

Janse Chris J., Franke-Fayard B, Mair GR, Ramesar J, Thiel C, Engelmann S, Matuschewski K, Gemert GJ van, Sauerwein RW, Waters AP. 2006. High efficiency transfection of Plasmodium berghei facilitates novel selection procedures. *Mol Biochem Parasitol* **145**:60–70. doi:10.1016/j.molbiopara.2005.09.007

Janse Chris J, Ramesar J, Waters AP. 2006. High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite Plasmodium berghei. *Nat Protoc* **1**:346–356. doi:10.1038/nprot.2006.53

Jenner RA, Wills MA. 2007. Opinion: The choice of model organisms in evo-devo. *Nat Rev Genet* **8**:311–319. doi:10.1038/nrg2062

Jensen LJ, Jensen TS, De Lichtenberg U, Brunak S, Bork P. 2006. Co-evolution of transcriptional and post-translational cell-cycle regulation. *Nature* **443**:594–597. doi:10.1038/nature05186

Jopling C, Sleep E, Raya M, Martí M, Raya A, Belmonte JCI. 2010. Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* **464**:606–609. doi:10.1038/nature08899

Jordan M, Schallhorn A, Wurm FM. 1996. Transfecting mammalian cells: Optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic Acids Res* **24**:596–601. doi:10.1093/nar/24.4.596

Jøstensen JP, Sperstad S, Johansen S, Landfald B. 2002. Molecular-phylogenetic, structural and biochemical features of a cold-adapted, marine ichthyosporean near the animal-fungal divergence, described from in vitro cultures. *Eur J Protistol*. doi:10.1078/0932-4739-00855

Joubes J, Chevalier C. 2000. Endoreduplication in higher plants. *Plant Mol Biol* **43**:735–745.

Kalucka J, Missiaen R, Georgiadou M, Schoors S, Lange C, De Bock K, Dewerchin M, Carmeliet P. 2015. Metabolic control of the cell cycle. *Cell Cycle*. doi:10.1080/15384101.2015.1090068

Karlsson J, Kroneis T, Jonasson E, Larsson E, Ståhlberg A. 2017. Transcriptomic Characterization of the Human Cell Cycle in Individual Unsynchronized Cells. *J Mol Biol* **429**:3909–3924. doi:10.1016/j.jmb.2017.10.011

Kawai S, Hashimoto W, Murata K. 2010. Transformation of Saccharomyces cerevisiae and other fungi. *Bioeng Bugs* **1**:395–403. doi:10.4161/bbug.1.6.13257

Kearsey SE, Cotterill S. 2003. Enigmatic variations: Divergent modes of regulating eukaryotic DNA replication. *Mol Cell*. doi:10.1016/S1097-2765(03)00441-6

Kelliher CM, Leman AR, Sierra CS, Haase SB. 2016. Investigating Conservation of the Cell-Cycle-Regulated Transcriptional Program in the Fungal Pathogen, Cryptococcus neoformans. *PLoS Genet* **12**:1–23. doi:10.1371/journal.pgen.1006453

Kilmartin J V. 2014. Lessons from yeast: The spindle pole body and the centrosome. *Philos Trans R Soc B Biol Sci* **369**. doi:10.1098/rstb.2013.0456

King N. 2004. The unicellular ancestry of animal development. *Dev Cell* **7**:313–325.

doi:10.1016/j.devcel.2004.08.010

King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I, Marr M, Pincus D, Putnam N, Rokas A, Wright KJ, Zuzow R, Dirks W, Good M, Goodstein D, Lemons D, Li W, Lyons JB, Morris A, Nichols S, Richter DJ, Salamov A, Sequencing JGI, Bork P, Lim WA, Manning G, Miller WT, McGinnis W, Shapiro H, Tjian R, Grigoriev I V, Rokhsar D. 2008. The genome of the choanoflagellate Monosiga brevicollis and the origin of metazoans. *Nature* **451**:783–8. doi:10.1038/nature06617

Kipreos ET, van den Heuvel S. 2019. Developmental control of the cell cycle: Insights from Caenorhabditis elegans. *Genetics* **211**:797–829. doi:10.1534/genetics.118.301643

Klar AJS. 2013. *Schizosaccharomyces japonicus* Yeast Poised to Become a Favorite Experimental Organism for Eukaryotic Research. *G3: Genes*|*Genomes*|*Genetics* **3**:1869–1873. doi:10.1534/g3.113.007187

Klein EA, Assoian RK. 2008. Transcriptional regulation of the cyclin D1 gene at a glance. *J Cell Sci* **121**:3853–3857. doi:10.1242/jcs.039131

Knoblich JA. 2008. Mechanisms of Asymmetric Stem Cell Division. *Cell*. doi:10.1016/j.cell.2008.02.007

Knoll AH. 2011. The Multiple Origins of Complex Multicellularity. *Annu Rev Earth Planet Sci*. doi:10.1146/annurev.earth.031208.100209

Kobayashi K, Suzuki T, Iwata E, Magyar Z, Bögre L, Ito M. 2015a. MYB3Rs, plant homologs of Myb oncoproteins, control cell cycle-regulated transcription and form DREAM-like complexes. *Transcription* **6**:106–111. doi:10.1080/21541264.2015.1109746

Kobayashi K, Suzuki Toshiya, Iwata E, Nakamichi N, Suzuki Takamasa, Chen P, Ohtani M, Ishida T, Hosoya H, Müller S, Leviczky T, Pettkó-Szandtner A, Darula Z, Iwamoto A, Nomoto M, Tada Y, Higashiyama T, Demura T, Doonan JH, Hauser M, Sugimoto K, Umeda M, Magyar Z, Bögre L, Ito M. 2015b. Transcriptional repression by MYB 3R proteins regulates plant organ growth . *EMBO J* **34**:1992–2007. doi:10.15252/embj.201490899

Koç A, Wheeler LJ, Mathews CK, Merrill GF. 2004. Hydroxyurea Arrests DNA Replication by a Mechanism that Preserves Basal dNTP Pools. *J Biol Chem* **279**:223–230. doi:10.1074/jbc.M303952200

Kotani S, Tugendreich S, Fujii M, Jorgensen PM, Watanabe N, Hoog C, Hieter P, Todokoro K. 1998. PKA and MPF-activated polo-like kinase regulate anaphase-promoting complex activity and mitosis progression. *Mol Cell*. doi:10.1016/S1097-2765(00)80037-4

Kronja I, Orr-Weaver TL. 2011. Translational regulation of the cell cycle: when, where, how and why? *Philos Trans R Soc B Biol Sci* **366**:3638–3652. doi:10.1098/rstb.2011.0084

Kuiper K. 2010. The Britannica guide to theories and ideas that changed the modern world, Turning points in history.

Landry JJM, Pyl PT, Rausch T, Zichner T, Tekkedil MM, Stütz AM, Jauch A, Aiyar RS, Pau G, Delhomme N, Gagneur J, Korbel JO, Huber W, Steinmetz LM. 2013. The genomic and transcriptomic landscape of a hela cell line. *G3 Genes, Genomes, Genet* **3**:1213–1224. doi:10.1534/g3.113.005777

Lang BF, O'Kelly C, Nerad T, Gray MW, Burger G. 2002. The closest unicellular relatives of animals. *Curr Biol*. doi:10.1016/S0960-9822(02)01187-9

Laoukili J, Kooistra MRH, Brás A, Kauw J, Kerkhoven RM, Morrison A, Clevers H, Medema RH. 2005. FoxM1 is required for execution of the mitotic programme and chromosome stability. *Nat Cell Biol*. doi:10.1038/ncb1217

Laumer CE, Fernández R, Lemer S, Combosch D, Kocot KM, Riesgo A, Andrade SCS, Sterrer W, Sørensen M V., Giribet G. 2019. Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proc R Soc B Biol Sci* **286**:20190831. doi:10.1098/rspb.2019.0831

Le Cam L. 1999. Timing of cyclin E gene expression depends on the regulated association of a bipartite repressor element with a novel E2F complex. *EMBO J* **18**:1878–1890. doi:10.1093/emboj/18.7.1878

Leadbeater BSC. 2015. The Choanoflagellates. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139051125

Lee MG, Nurse P. 1987. Complementation used to clone a human homologue of the fission yeast cell cycle control gene cdc2. *Nature* **327**:31–35. doi:10.1038/327031a0

Leonelli S, Ankeny RA. 2013. What makes a model organism? *Endeavour* **37**:209–212. doi:10.1016/j.endeavour.2013.06.001

Leonelli S, Ankeny RA. 2012. Re-thinking organisms: The impact of databases on model organism biology. *Stud Hist Philos Sci Part C Stud Hist Philos Biol Biomed Sci* **43**:29–36. doi:10.1016/j.shpsc.2011.10.003

Lerche K, Hallmann A. 2014. Stable nuclear transformation of Pandorina morum. *BMC Biotechnol* **14**:65. doi:10.1186/1472-6750-14-65

Lerche K, Hallmann A. 2013. Stable nuclear transformation of Eudorina elegans. *BMC Biotechnol* **13**:11. doi:10.1186/1472-6750-13-11

Lerche K, Hallmann A. 2009. Stable nuclear transformation of Gonium pectorale. *BMC Biotechnol* **9**:64. doi:10.1186/1472-6750-9-64

Levin TC, Greaney AJ, Wetzel L, King N. 2014. The Rosetteless gene controls development in the choanoflagellate S. rosetta. *Elife*. doi:10.7554/eLife.04070

Levin TC, King N. 2013. Evidence for sex and recombination in the choanoflagellate Salpingoeca rosetta. *Curr Biol*.

Li Z. 2012. Regulation of the Cell Division Cycle in Trypanosoma brucei. *Eukaryot Cell* **11**:1180–1190. doi:10.1128/ec.00145-12

Li Z, Wang CC. 2003. A PHO80-like cyclin and a B-type cyclin control the cell cycle of the procyclic form of Trypanosoma brucei. *J Biol Chem* **278**:20652–20658. doi:10.1074/jbc.M301635200

Libby E, Rainey PB. 2011. Exclusion rules, bottlenecks and the evolution of stochastic phenotype switching. *Proc R Soc B Biol Sci*. doi:10.1098/rspb.2011.0146

Lim S, Kaldis P. 2013. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development* **140**:3079–3093. doi:10.1242/dev.091744

Lindås AC, Bernander R. 2013. The cell cycle of archaea. *Nat Rev Microbiol*. doi:10.1038/nrmicro3077

Lindegren CC. 1932. The Genetics of Neurospora-II. Segregation of the Sex Factors in Asci of N. crassa, N. sitophila, and N. tetrasperma. *Bull Torrey Bot Club*. doi:10.2307/2480466

Litovchick L, Sadasivam S, Florens L, Zhu X, Swanson SK, Velmurugan S, Chen R, Washburn MP, Liu XS, DeCaprio JA. 2007. Evolutionarily Conserved Multisubunit RBL2/p130 and E2F4 Protein Complex Represses Human Cell Cycle-Dependent Genes in Quiescence. *Mol Cell*. doi:10.1016/j.molcel.2007.04.015

Liu J, Kipreos ET. 2000. Evolution of Cyclin-Dependent Kinases (CDKs) and CDK-Activating Kinases (CAKs): Differential Conservation of CAKs in Yeast and Metazoa. *Mol Biol Evol* **17**:1061–1074. doi:10.1093/oxfordjournals.molbev.a026387

Liu Y, Beyer A, Aebersold R. 2016. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**:535–550. doi:10.1016/j.cell.2016.03.014

Lodhi N, Kossenkov A V., Tulin A V. 2014. Bookmarking promoters in mitotic chromatin: Poly(ADP-ribose)polymerase-1 as an epigenetic mark. *Nucleic Acids Res* **42**:7028–7038. doi:10.1093/nar/gku415

Lohka MJ. 1989. Mitotic control by metaphase-promoting factor and cdc proteins. *J Cell Sci* **92 (Pt 2)**:131–135.

López-Escardó D, López-García P, Moreira D, Ruiz-Trillo I, Torruella G. 2018. Parvularia atlantis gen. et sp. nov., a Nucleariid Filose Amoeba (Holomycota, Opisthokonta). *J Eukaryot Microbiol* **65**:170–179. doi:10.1111/jeu.12450

Loyter A, Scangos GA, Ruddle FH. 1982. Mechanisms of DNA uptake by mammalian cells: Fate of exogenously added DNA monitored by the use of fluorescent dyes. *Proc Natl Acad Sci U S A* **79**:422–426. doi:10.1073/pnas.79.2.422

Lukeš J, Archibald JM, Keeling PJ, Doolittle WF, Gray MW. 2011. How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life* **63**:528–537. doi:10.1002/iub.489

Lundgren M, Bernander R. 2007. Genome-wide transcription map of an archaeal cell cycle. *Proc Natl Acad Sci U S A*. doi:10.1073/pnas.0611333104

Madigan M, Martinko J, Dunlap P, Clark D. 2008. Brock Biology of microorganisms 12th edn. *Int Microbiol*.

Marshall WL, Celio G, McLaughlin DJ, Berbee ML. 2008. Multiple Isolations of a Culturable, Motile Ichthyosporean (Mesomycetozoa, Opisthokonta), Creolimax fragrantissima n. gen., n. sp., from Marine Invertebrate Digestive Tracts. *Protist*. doi:10.1016/j.protis.2008.03.003

Marston AL, Amon A. 2004. Meiosis: Cell-cycle controls shuffle and deal. *Nat Rev Mol Cell Biol* **5**:983–997. doi:10.1038/nrm1526

Martens H, Novotny J, Oberstrass J, Steck TL, Postlethwait P, Nellen W. 2002. RNAi in Dictyostelium: The role of RNA-directed RNA polymerases and double-stranded RNase. *Mol Biol Cell* **13**:445–453. doi:10.1091/mbc.01-04-0211

Martín-Castellanos C, Blanco MA, De Prada JM, Moreno S. 2000. The puc1 cyclin regulates the G1 phase of the fission yeast cell cycle in response to cell size. *Mol Biol Cell* **11**:543–554. doi:10.1091/mbc.11.2.543

Matsushime H, Quelle DE, Shurtleff SA, Shibuya M, Sherr CJ, Kato JY. 1994. D-type cyclin-dependent kinase activity in mammalian cells. *Mol Cell Biol*. doi:10.1128/mcb.14.3.2066

Matsushime H, Roussel MF, Ashmun RA, Sherr CJ. 1991. Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle. *Cell* **65**:701–713. doi:10.1016/0092-8674(91)90101-4

Matthews H, Duffy CW, Merrick CJ. 2018. Checks and balances? DNA replication and the cell cycle in Plasmodium, Parasites & Vectors. BioMed Central Ltd. doi:10.1186/s13071-018-2800-1

Measday V, Moore L, Retnakaran R, Lee J, Donoviel M, Neiman AM, Andrews B. 1997. A family of cyclin-like proteins that interact with the Pho85 cyclin-dependent kinase. *Mol Cell Biol* **17**:1212–1223. doi:10.1128/mcb.17.3.1212

Medina EM, Turner JJ, Gordân R, Skotheim JM, Buchler NE. 2016. Punctuated evolution and transitional hybrid network in an ancestral cell cycle of fungi. *Elife* **5**:1–23. doi:10.7554/eLife.09492.001

Mendenhall MD, Hodge a E. 1998. Regulation of CDC28 cyclin-dependent protein kinase activity during the cell cycle of the yeast *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev*.

Mendoza L, Taylor JW, Ajello L. 2002. The Class Mesomycetozoea: A Heterogeneous Group of Microorganisms at the Animal-Fungal Boundary. *Annu Rev Microbiol*. doi:10.1146/annurev.micro.56.012302.160950

Menges M, Hennig L, Gruissem W, Murray JAH. 2003. Genome-wide gene expression in an Arabidopsis cell suspension. *Plant Mol Biol* **53**:423–442. doi:10.1023/B:PLAN.0000019059.56489.ca

Menges M, Murray JAH. 2002. Synchronous Arabidopsis suspension cultures for analysis of cell-cycle gene activity. *Plant J* **30**:203–212. doi:10.1046/j.1365-313X.2002.01274.x

Mikhailov K V., Konstantinova A V., Nikitin MA, Troshin P V., Rusin LY, Lyubetsky VA, Panchin Y V., Mylnikov AP, Moroz LL, Kumar S, Aleoshin V V. 2009. The origin of Metazoa: A transition from temporal to spatial cell differentiation. *BioEssays* **31**:758–768. doi:10.1002/bies.200800214

Mitchison JM, Creanor J. 1971. Induction synchrony in the fission yeast Schizosaccharomyces pombe. *Exp Cell Res*. doi:10.1016/0014-4827(71)90421-6

Mondesert O, McGowan CH, Russell P. 2015. Cig2, a B-type cyclin, promotes the onset of S in Schizosaccharomyces pombe. *Mol Cell Biol* **16**:1527–1533. doi:10.1128/mcb.16.4.1527

Mora C, Tittensor DP, Adl S, Simpson AGB, Worm B. 2011. How many species are there on earth and in the ocean? *PLoS Biol*. doi:10.1371/journal.pbio.1001127

Morgan DO. 2007. Cell Cycle: Principles of Control. *Yale J Biol Med*. doi:10.1093/icb/icm066

Morgan, Sturtevant, Muller, Bridges. 1915. Mechanism of Mendelian Heridity. *Trans Am Microsc Soc* **34**:293. doi:10.2307/3221480

Morse D, Daoust P, Benribague S. 2016. A Transcriptome-based Perspective of Cell Cycle Regulation in Dinoflagellates. *Protist* **167**:610–621.

doi:10.1016/j.protis.2016.10.002

Moshe Y, Boulaire J, Pagano M, Hershko A. 2004. Role Polo-like kinase in the degradation of early mitotic inhibitor 1, a regulator of the anaphase promoting complex/cyclosome. *Proc Natl Acad Sci U S A* **101**:7937–7942. doi:10.1073/pnas.0402442101

Müller B, Grossniklaus U. 2010. Model organisms - A historical perspective. *J Proteomics* **73**:2054–2063. doi:10.1016/j.jprot.2010.08.002

Murray AW, Kirschner MW. 1989. Cyclin synthesis drives the early embryonic cell cycle. *Nature* **339**:275–280. doi:10.1038/339275a0

Musa J, Aynaud MM, Mirabeau O, Delattre O, Grünewald TG. 2017. MYBL2 (B-Myb): a central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis. *Cell Death Dis* **8**:e2895. doi:10.1038/cddis.2017.244

Nacu E, Tanaka EM. 2011. Limb Regeneration: A New Development? *Annu Rev Cell Dev Biol* **27**:409–440. doi:10.1146/annurev-cellbio-092910-154115

Nagano T, Hashimoto T, Nakashima A, Hisanaga S, Kikkawa U, Kamada S. 2013. Cyclin I is involved in the regulation of cell cycle progression. *Cell Cycle*. doi:10.4161/cc.25623

Nagy LG, Kovács GM, Krizsán K. 2018. Complex multicellularity in fungi: evolutionary convergence, single origin, or both? *Biol Rev* **93**:1778–1794. doi:10.1111/brv.12418

Nasmyth K. 1995. Evolution of the cell cycle. *Philos Trans R Soc Lond B Biol Sci* **349**:271–281. doi:10.1098/rstb.1995.0113

Nasmyth K, Dirick L. 1991. The role of SWI4 and SWI6 in the activity of G1 cyclins in yeast. *Cell* **66**:995–1013. doi:10.1016/0092-8674(91)90444-4

Naumov A, Kratzer S, Ting LM, Kim K, Suvorova ES, White MW. 2017. The Toxoplasma centrocone houses cell cycle regulatory factors. *MBio* **8**. doi:10.1128/mBio.00579-17

Newcomb LL, Diderich JA, Slattery MG, Heideman W. 2003. Glucose regulation of Saccharomyces cerevisiae cell cycle genes. *Eukaryot Cell*.

doi:10.1128/EC.2.1.143-149.2003

Nguyen TA, Cissé OH, Yun Wong J, Zheng P, Hewitt D, Nowrousian M, Stajich JE, Jedd G. 2017. Innovation and constraint leading to complex multicellularity in the Ascomycota. *Nat Commun* **8**:14444. doi:10.1038/ncomms14444

Novoa I, Gallego J, Ferreira PG, Mendez R. 2010. Mitotic cell-cycle progression is regulated by CPEB1 and CPEB4-dependent translational control. *Nat Cell Biol* **12**:447–456. doi:10.1038/ncb2046

Nüsslein-volhard C, Wieschaus E. 1980. Mutations affecting segment number and polarity in drosophila. *Nature* **287**:795–801. doi:10.1038/287795a0

Ohtani K, Degregori J, Nevins JR. 1995. Regulation of the cyclin E gene by transcription factor E2F1. *Proc Natl Acad Sci U S A* **92**:12146–12150. doi:10.1073/pnas.92.26.12146

Olson RJ, Chisholm SW. 1986. Effects of light and nitrogen limitation on the cell cycle of the dinoflagellate Amphidinium carteri. *J Plankton Res* **8**:785–793. doi:10.1093/plankt/8.4.785

Ondracka A, Dudin O, Ruiz-Trillo I. 2018. Decoupling of Nuclear Division Cycles and Cell Size during the Coenocytic Growth of the Ichthyosporean Sphaeroforma arctica. *Curr Biol*. doi:10.1016/j.cub.2018.04.074

Orlando DA, Lin CY, Bernard A, Wang JY, Socolar JES, Iversen ES, Hartemink AJ, Haase SB. 2008. Global control of cell-cycle transcription by coupled CDK and network oscillators. *Nature* **453**:944–947. doi:10.1038/nature06955

Owczarzak A, Stibbs HH, Bayne CJ. 1980. The destruction of Schistosoma mansoni mother sporocysts in vitro by amoebae isolated from Biomphalaria glabrata: an ultrastructural study. *J Invertebr Pathol* **35**:26–33. doi:10.1016/0022-2011(80)90079-8

Pang KM, Lynes M a, Knecht D a. 1999. Variables controlling the expression level of exogenous genes in Dictyostelium. *Plasmid* **41**:187–97. doi:10.1006/plas.1999.1391

Paps J, Holland PWH. 2018. Reconstruction of the ancestral metazoan genome reveals

an increase in genomic novelty. *Nat Commun* **9**:1–8. doi:10.1038/s41467-018-04136-5

Peeper DS, Parker LL, Ewen ME, Toebes M, Hall FL, Xu M, Zantema A, van der Eb AJ, Piwnica-Worms H. 1993. A- and B-type cyclins differentially modulate substrate specificity of cyclin-cdk complexes. *EMBO J* **12**:1947–1954. doi:10.1002/j.1460-2075.1993.tb05844.x

Peña-Diaz J, Hegre SA, Anderssen E, Aas PA, Mjelle R, Gilfillan GD, Lyle R, Drabløs F, Krokan HE, Sætrom P. 2013. Transcription profiling during the cell cycle shows that a subset of Polycomb-targeted genes is upregulated during DNA replication. *Nucleic Acids Res* **41**:2846–2856. doi:10.1093/nar/gks1336

Pentz JT, Limberg T, Beermann N, Ratcliff WC. 2015. Predator Escape: An Ecologically Realistic Scenario for the Evolutionary Origins of Multicellularity. *Evol Educ Outreach* **8**. doi:10.1186/s12052-015-0041-8

Peter M, Nakagawa J, Dorée M, Labbé JC, Nigg EA. 1990. In Vitro Disassembly of the Nuclear Lamina and M Phase-Specific Phosphorylation of Lamins by cdc2 Kinase. *Cell* **61**:591–602.

Philpott A, Yew PR. 2008. The xenopus cell cycle: An overview. *Mol Biotechnol* **39**:9–19. doi:10.1007/s12033-008-9033-z

Polanowska J, Fabbrizio E, Le Cam L, Trouche D, Emiliani S, Herrera R, Sardet C. 2001. The periodic down regulation of Cyclin E gene expression from exit of mitosis to end of G1 is controlled by a deacetylase- and E2F-associated bipartite repressor element. *Oncogene* **20**:4115–4127. doi:10.1038/sj.onc.1204514

Polyak K, Kato JY, Solomon MJ, Sherr CJ, Massague J, Roberts JM, Koff A. 1994. p27(Kip1), a cyclin-Cdk inhibitor, links transforming growth factor-β and contact inhibition to cell cycle arrest. *Genes Dev* **8**:9–22.

Pomerantz RT, O'Donnell M. 2010. Direct restart of a replication fork stalled by a head-on RNA polymerase. *Science* . doi:10.1126/science.1179595

Pomerening JR, Sontag ED, Ferrell JE. 2003. Building a cell cycle oscillator: Hysteresis and bistability in the activation of Cdc2. *Nat Cell Biol*. doi:10.1038/ncb954

Pomerening JR, Sun YK, Ferrell JE. 2005. Systems-level dissection of the cell-cycle oscillator: Bypassing positive feedback produces damped oscillations. *Cell* **122**:565–578. doi:10.1016/j.cell.2005.06.016

Pop R, Shearstone JR, Shen Q, Liu Y, Hallstrom K, Koulnis M, Gribnau J, Socolovsky M. 2010. A key commitment step in erythropoiesis is synchronized with the cell cycle clock through mutual inhibition between PU.1 and S-phase progression. *PLoS Biol* **8**. doi:10.1371/journal.pbio.1000484

Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK, Thurman RE, Cheng Y, Gülsoy G, Dennis JH, Snyder MP, Stamatoyannopoulos JA, Taylor J, Hardison RC, Kahveci T, Ren B, Gilbert DM. 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature*. doi:10.1038/nature13986

Poss KD, Wilson LG, Keating MT. 2002. Heart regeneration in zebrafish. *Science* **298**:2188–2190. doi:10.1126/science.1077857

Pozarowski P, Darzynkiewicz Z. 2005. Analysis of Cell Cycle by Flow CytometryCheckpoint Controls and Cancer. New Jersey: Humana Press. pp. 301–312. doi:10.1385/1-59259-811-0:301

Pramila T, Wu W, Miles S, Noble WS, Breeden LL. 2006. The Forkhead transcription factor Hcm1 regulates chromosome segregation genes and fills the S-phase gap in the transcriptional circuitry of the cell cycle. *Genes Dev*. doi:10.1101/gad.1450606

Prelich G. 2012. Gene Overexpression: Uses, Mechanisms, and Interpretation. *Genetics* **190**:841–854. doi:10.1534/genetics.111.136911

Proserpio V, Casale FP, Natarajan KN, Marioni JC, Stegle O, Teichmann SA, Scialdone A, Buettner F, Theis FJ. 2015. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol* **33**:155–160. doi:10.1038/nbt.3102

Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, Terry A, Shapiro H, Lindquist E, Kapitonov V V., Jurka J, Genikhovich G, Grigoriev I V., Lucas SM, Steele RE, Finnerty JR, Technau U, Martindale MQ, Rokhsar DS. 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic

organization. *Science* **317**:86–94. doi:10.1126/science.1139158

Raghu-kumar S. 1987. Occurrence of the Thraustochytrid, Corallochytrium limacisporum gen. et sp. nov. in the Coral Reef Lagoons of the Lakshadweep Islands in the Arabian Sea. *Bot Mar*. doi:10.1515/botm.1987.30.1.83

Rainey PB. 2007. Unity from conflict. *Nature*. doi:10.1038/446616a

Rainey PB, Kerr B. 2010. Cheats as first propagules: A new hypothesis for the evolution of individuality during the transition from single cells to multicellularity. *BioEssays*. doi:10.1002/bies.201000039

Rajan I, Rabindran R, Nithya N, Lakshmipriya T, Jayasree PR, Kumar PRM. 2014. Assessment of cell cycle phase-specific effects of zerumbone on mitotically synchronous surface cultures of Physarum polycephalum. *Protoplasma* **251**:931–941. doi:10.1007/s00709-013-0605-9

Ran FA, Hsu PPD, Wright J, Agarwala V, Scott D a, Zhang F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**:2281–308. doi:10.1038/nprot.2013.143

Rao SS, Chu C, Kohtz DS. 1994. Ectopic expression of cyclin D1 prevents activation of gene transcription by myogenic basic helix-loop-helix regulators. *Mol Cell Biol* **14**:5259–5267. doi:10.1128/mcb.14.8.5259

Ratcliff WC, Fankhauser JD, Rogers DW, Greig D, Travisano M. 2015. Origins of multicellular evolvability in snowflake yeast. *Nat Commun* **6**:1–9. doi:10.1038/ncomms7102

Ratcliff WC, Herron MD, Howell K, Pentz JT, Rosenzweig F, Travisano M. 2013. Experimental evolution of an alternating uni- and multicellular life cycle in Chlamydomonas reinhardtii. *Nat Commun* **4**:2742. doi:10.1038/ncomms3742

Reed SI. 1980. The selection of S. cerevisiae mutants defective in the start event of cell division. *Genetics* **95**:561–577.

Rejlová L, Chrtek J, Trávníček P, Lučanová M, Vít P, Urfus T. 2019. Polyploid evolution: The ultimate way to grasp the nettle. *PLoS One* **14**:e0218389. doi:10.1371/journal.pone.0218389

Reynolds D, Shi BJ, McLean C, Katsis F, Kemp B, Dalton S. 2003. Recruitment of Thr 319-phosphorylated Ndd1p to the FHA domain of Fkh2p requires C1b kinase activity: A mechanism for CLB cluster gene activation. *Genes Dev* **17**:1789–1802. doi:10.1101/gad.1074103

Rhind N, Chen Z, Yassour M, Thompson DA, Haas BJ, Habib N, Wapinski I, Roy S, Lin MF, Heiman DI, Young SK, Furuya K, Guo Y, Pidoux A, Chen HM, Robbertse B, Goldberg JM, Aoki K, Bayne EH, Berlin AM, Desjardins CA, Dobbs E, Dukaj L, Fan L, FitzGerald MG, French C, Gujja S, Hansen K, Keifenheim D, Levin JZ, Mosher RA, Müller CA, Pfiffner J, Priest M, Russ C, Smialowska A, Swoboda P, Sykes SM, Vaughn M, Vengrova S, Yoder R, Zeng Q, Allshire R, Baulcombe D, Birren BW, Brown W, Ekwall K, Kellis M, Leatherwood J, Levin H, Margalit H, Martienssen R, Nieduszynski CA, Spatafora JW, Friedman N, Dalgaard JZ, Baumann P, Niki H, Regev A, Nusbaum C. 2011. Comparative functional genomics of the fission yeasts. *Science* **332**:930–936. doi:10.1126/science.1203357

Rhind N, Russell P. 2000. Chk1 and Cds1: linchpins of the DNA damage and replication checkpoint pathways. *J Cell Sci* **113 (Pt 2**:3889–96.

Richardson HE, Wittenberg C, Cross F, Reed SI. 1989. An essential G1 function for cyclin-like proteins in yeast. *Cell* **59**:1127–1133. doi:10.1016/0092-8674(89)90768-X

Richter DJ, Fozouni P, Eisen MB, King N. 2018. Gene family innovation, conservation and loss on the animal stem lineage. *Elife*. doi:10.7554/eLife.34226

Richter DJ, King N. 2013. The Genomic and Cellular Foundations of Animal Origins. *Annu Rev Genet*. doi:10.1146/annurev-genet-111212-133456

Robertson ADJ, Grutsch JF. 1981. Aggregation in Dictyostelium discoideum. *Cell*. doi:10.1016/0092-8674(81)90087-8

Rokas A. 2008. The molecular origins of multicellular transitions. *Curr Opin Genet Dev* **18**:472–478. doi:10.1016/j.gde.2008.09.004

Ruiz-Trillo I, Burger G, Holland PWH, King N, Lang BF, Roger AJ, Gray MW. 2007. The origins of multicellularity: a multi-taxon genome initiative. *Trends Genet* **23**:113–118. doi:10.1016/j.tig.2007.01.005

Ruiz-Trillo I, Inagaki Y, Davis LA, Sperstad S, Landfald B, Roger AJ. 2004. Capsaspora owczarzaki is an independent opisthokont lineage. *Curr Biol* **14**:R946–R947. doi:10.1016/j.cub.2004.10.037

Ruiz-Trillo I, Roger AJ, Burger G, Gray MW, Lang BF. 2008. A phylogenomic investigation into the origin of Metazoa. *Mol Biol Evol* **25**:664–672. doi:10.1093/molbev/msn006

Rusch HP, Sachsenmaier W, Behrens K, Gruter V. 1966. Synchronization of mitosis by the fusion of the plasmodia of Physarum polycephalum. *J Cell Biol*. doi:10.1083/jcb.31.1.204

Russell JJ, Theriot JA, Sood P, Marshall WF, Landweber LF, Fritz-Laylin L, Polka JK, Oliferenko S, Gerbich T, Gladfelter A, Umen J, Bezanilla M, Lancaster MA, He S, Gibson MC, Goldstein B, Tanaka EM, Hu CK, Brunet A. 2017. Non-model model organisms, BMC Biology. BMC Biology. doi:10.1186/s12915-017-0391-5

Rustici G, Mata J, Kivinen K, Lió P, Penkett CJ, Burns G, Hayles J, Brazma A, Nurse P, Bähler J. 2004. Periodic gene expression program of the fission yeast cell cycle. *Nat Genet* **36**:809–817. doi:10.1038/ng1377

Sachsenmaier W, Remy U, Plattner-Schobel R. 1972. Initiation of synchronous mitosis in Physarum polycephalum. A model of the control of cell division in eukariots. *Exp Cell Res* **73**:41–48. doi:10.1016/0014-4827(72)90099-7

Sadasivam S, Duan S, DeCaprio JA. 2012. The MuvB complex sequentially recruits B-Myb and FoxM1 to promote mitotic gene expression. *Genes Dev* **26**:474–489. doi:10.1101/gad.181933.111

Samerier M. 2011. Functional analyses of microtubule and centrosome-associated proteins in Dictyostelium dicsoideum. University of Postdam.

Santamaría D, Barrière C, Cerqueira A, Hunt S, Tardy C, Newton K, Cáceres JF, Dubus P, Malumbres M, Barbacid M. 2007. Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* **448**:811–815. doi:10.1038/nature06046

Schiestl RH, Gietz RD. 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* **16**:339–346. doi:10.1007/BF00340712

Schmerler S, Wessel GM. 2011. Polar bodies-more a lack of understanding than a lack of respect. *Mol Reprod Dev* **78**:3–8. doi:10.1002/mrd.21266

Schmid V, Alder H, Plickert G, Weber C. 1988. Transdifferentiation from striated muscle of medusae in vitro. *Cell Differ Dev* **25**:137–146. doi:10.1016/0922-3371(88)90110-4

Schmoller KM, Turner JJ, Kõivomägi M, Skotheim JM. 2015. Dilution of the cell cycle inhibitor Whi5 controls budding-yeast cell size. *Nature* **526**:268–272. doi:10.1038/nature14908

Schwob E, Nasmyth K. 1993. CLB5 and CLB6, a new pair of B cyclins involved in DNA replication in Saccharomyces cerevisiae. *Genes Dev* **7**:1160–1175. doi:10.1101/gad.7.7a.1160

Scialdone A, Natarajan KN, Saraiva LR, Proserpio V, Teichmann SA, Stegle O, Marioni JC, Buettner F. 2015. Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods* **85**:54–61. doi:10.1016/j.ymeth.2015.06.021

Sebé-Pedrós A, Ballaré C, Parra-Acero H, Chiva C, Tena JJ, Sabidó E, Gómez-skarmeta JL, Di Croce L, Ruiz-Trillo II. 2016a. The Dynamic Regulatory Genome of Capsaspora and the Origin of Animal Multicellularity. *Cell* **165**:1224–1237. doi:10.1016/j.cell.2016.03.034

Sebé-Pedrós A, Degnan BM, Ruiz-Trillo I. 2017. The origin of Metazoa, a unicellular perspective. *Nat Rev Genet* **in press**. doi:10.1038/nrg.2017.21

Sebé-Pedrós A, Irimia M, del Campo J, Parra-Acero H, Russ C, Nusbaum C, Blencowe BJ, Ruiz-Trillo I. 2013. Regulated aggregative multicellularity in a close unicellular relative of metazoa. *Elife* **2**:1–20. doi:10.7554/eLife.01287

Sebé-Pedrós A, Peña MI, Capella-Gutiérrez S, Antó M, Gabaldón T, Ruiz-Trillo II, Sabidó E, Gabaldon T, Ruiz-Trillo II, Sabidó E. 2016b. High-Throughput Proteomics Reveals the Unicellular Roots of Animal Phosphosignaling and Cell Differentiation. *Dev Cell* **39**:186–197. doi:10.1016/j.devcel.2016.09.019

Sebé-Pedrós A, Roger AJ, Lang FB, King N, Ruiz-Trillo I. 2010. Ancient origin of the integrin-mediated adhesion and signaling machinery. *Proc Natl Acad Sci* **107**:10142–10147. doi:10.1073/pnas.1002257107

Shalchian-Tabrizi K, Minge MA, Espelund M, Orr R, Ruden T, Jakobsen KS, Cavalier-Smith T. 2008. Multigene Phylogeny of Choanozoa and the Origin of Animals. *PLoS One* **3**:e2098. doi:10.1371/journal.pone.0002098

Shepard JL, Stern HM, Pfaff KL, Amatruda JF. 2004. Analysis of the cell cycle in zebrafish embryos. *Methods Cell Biol* **2004**:109–125.

Sherr CJ. 1995. D-type cyclins. *Trends Biochem Sci* **20**:187–190. doi:10.1016/S0968-0004(00)89005-2

Shi X, Ma M, Lin S. 2017. Cell cycle-dependent expression dynamics of G1/S specific cyclin, cellulose synthase and cellulase in the dinoflagellate Prorocentrum donghaiense. *Front Microbiol* **8**. doi:10.3389/fmicb.2017.01118

Shirayama M, Tóth A, Gálová M, Nasmyth K. 1999. APC(Cdc20) promotes exit from mitosis by destroying the anaphase inhibitor Pds1 and cyclin Clb5. *Nature*. doi:10.1038/46080

Singh AM, Chappell J, Trost R, Lin L, Wang T, Tang J, Wu H, Zhao S, Jin P, Dalton S. 2013. Cell-cycle control of developmentally regulated transcription factors accounts for heterogeneity in human pluripotent cells. *Stem Cell Reports* **1**:532–544. doi:10.1016/j.stemcr.2013.10.009

Singh AM, Sun Y, Li L, Zhang W, Wu T, Zhao S, Qin Z, Dalton S. 2015. Cell-Cycle Control of Bivalent Epigenetic Domains Regulates the Exit from Pluripotency. *Stem Cell Reports* **5**:323–336. doi:10.1016/j.stemcr.2015.07.005

Sivan G, Elroy-Stein O. 2008. Regulation of mRNA Translation during cellular division. *Cell Cycle* **7**:741–744. doi:10.4161/cc.7.6.5596

Skapek SX, Rhee J, Spicer DB, Lassar AB. 1995. Inhibition of myogenic differentiation in proliferating myoblasts by cyclin D1-dependent kinase. *Science* **267**:1022–1024. doi:10.1126/science.7863328

Skotheim JM, Di Talia S, Siggia ED, Cross FR. 2008. Positive feedback of G1 cyclins ensures coherent cell cycle entry. *Nature* **454**:291–296. doi:10.1038/nature07118

Slater ML. 1973. Effect of reversible inhibition of deoxyribonucleic acid synthesis on the yeast cell cycle. *J Bacteriol*.

Soufi A, Dalton S. 2016. Cycling through developmental decisions: How cell cycle dynamics control pluripotency, differentiation and reprogramming. *Dev* **143**:4301–4311. doi:10.1242/dev.142075

Soufi A, Donahue G, Zaret KS. 2012. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**:994–1004. doi:10.1016/j.cell.2012.09.045

Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, Futcher B. 1998. Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast Saccharomyces cerevisiae by Microarray Hybridization. *Mol Biol Cell*. doi:10.1091/mbc.9.12.3273

Steenwyk JL, Opulente DA, Kominek J, Shen XX, Zhou X, Labella AL, Bradley NP, Eichman BF, Čadež N, Libkind D, De Virgilio J, Hulfachor AB, Kurtzman CP, Hittinger CT, Rokas A. 2019. Extensive loss of cell-cycle and DNA repair genes in an ancient lineage of bipolar budding yeasts. *PLoS Biol*. doi:10.1371/journal.pbio.3000255

Stibbs HH, Owczarzak A, Bayne CJ, DeWan P. 1979. Schistosome sporocyst-killing amoebae isolated from Biomphalaria glabrata. *J Invertebr Pathol* **33**:159–170. doi:10.1016/0022-2011(79)90149-6

Strasser K, Bloomfield G, MacWilliams A, Ceccarelli A, MacWilliams H, Tsang A. 2012. A retinoblastoma orthologue is a major regulator of S-phase, mitotic, and developmental gene expression in Dictyostelium. *PLoS One* **7**. doi:10.1371/journal.pone.0039914

Strickfaden SC, Winters MJ, Ben-Ari G, Lamson RE, Tyers M, Pryciak PM. 2007. A Mechanism for Cell-Cycle Regulation of MAP Kinase Signaling in a Yeast Differentiation Pathway. *Cell*. doi:10.1016/j.cell.2006.12.032

Stumpf CR, Moreno M V., Olshen AB, Taylor BS, Ruggero D. 2013. The Translational Landscape of the Mammalian Cell Cycle. *Mol Cell* **52**:574–582. doi:10.1016/j.molcel.2013.09.018

Sudo T, Ota Y, Kotani S, Nakao M, Takami Y, Takeda S, Saya H. 2001. Activation of Cdh1-dependent APC is required for G1 cell cycle arrest and DNA damage-induced

G2 checkpoint in vertebrate cells. *EMBO J* **20**:6499–6508. doi:10.1093/emboj/20.22.6499

Suga H, Chen Z, de Mendoza A, Sebé-Pedrós A, Brown MW, Kramer E, Carr M, Kerner P, Vervoort M, Sánchez-Pons N, Torruella G, Derelle R, Manning G, Lang BF, Russ C, Haas BJ, Roger AJ, Nusbaum C, Ruiz-Trillo I. 2013. The Capsaspora genome reveals a complex unicellular prehistory of animals. *Nat Commun* **4**:1–9. doi:10.1038/ncomms3325

Suga H, Dacre M, De Mendoza A, Shalchian-Tabrizi K, Manning G, Ruiz-Trillo I. 2012. Genomic survey of premetazoans shows deep conservation of cytoplasmic tyrosine kinases and multiple radiations of receptor tyrosine kinases. *Sci Signal*. doi:10.1126/scisignal.2002733

Suga H, Ruiz-Trillo I. 2013. Development of ichthyosporeans sheds light on the origin of metazoan multicellularity. *Dev Biol* **377**:284–292. doi:10.1016/j.ydbio.2013.01.009

Sugiyama M, Sakaue-Sawano A, Iimura T, Fukami K, Kitaguchi T, Kawakami K, Okamoto H, Higashijima SI, Miyawaki A. 2009. Illuminating cell-cycle progression in the developing zebrafish embryo. *Proc Natl Acad Sci U S A* **106**:20812–20817. doi:10.1073/pnas.0906464106

Sun M, Bernard LP, Dibona VL, Wu Q, Zhang H. 2013. Calcium phosphate transfection of primary hippocampal neurons. *J Vis Exp*.

Suvorova ES, Radke JB, Ting LM, Vinayak S, Alvarez CA, Kratzer S, Kim K, Striepen B, White MW. 2013. A nucleolar AAA-NTPase is required for parasite division. *Mol Microbiol*. doi:10.1111/mmi.12367

Szathmáry E. 2015. Toward major evolutionary transitions theory 2.0. *Proc Natl Acad Sci U S A* **112**. doi:10.1073/pnas.1421398112

Szathmary E, Smith JM. 1995. The major evolutionary transitions Transition from independent replicators. *Nature* **374**:227–232.

Tan ALC, Rida PCG, Surana U. 2005. Essential tension and constructive destruction: The spindle checkpoint and its regulatory links with mitotic exit. *Biochem J*. doi:10.1042/BJ20041415

Taroncher-Oldenburg G, Kulis DM, Anderson DM. 1999. Coupling of saxitoxin biosynthesis to the G1 phase of the cell cycle in the dinoflagellate Alexandrin fundyense: Temperature and nutrient effects. *Nat Toxins* **7**:207–219. doi:10.1002/1522-7189(200009/10)7:5<207::AID-NT61>3.0.CO;2-Q

Tekaia F. 2016. Inferring Orthologs: Open Questions and Perspectives. *Genomics Insights* **9**:GEI.S37925. doi:10.4137/GEI.S37925

Tiengwe C, Marcello L, Farr H, Gadelha C, Burchmore R, Barry JD, Bell SD, McCulloch R. 2012. Identification of ORC1/CDC6-interacting factors in trypanosoma brucei reveals critical features of origin recognition complex architecture. *PLoS One* **7**. doi:10.1371/journal.pone.0032674

Tong SM. 1997. Heterotrophic flagellates and other protists from Southampton Water, U.K. *Ophelia* **47**:71–131. doi:10.1080/00785236.1997.10427291

Torruella G, De Mendoza A, Grau-Bové X, Antó M, Chaplin MA, Del Campo J, Eme L, Pérez-Cordón G, Whipps CM, Nichols KM, Paley R, Roger AJ, Sitjà-Bobadilla A, Donachie S, Ruiz-Trillo I. 2015. Phylogenomics Reveals Convergent Evolution of Lifestyles in Close Relatives of Animals and Fungi. *Curr Biol* **25**:2404–2410. doi:10.1016/j.cub.2015.07.053

Torruella G, Derelle R, Paps J, Lang BF, Roger AJ, Shalchian-Tabrizi K, Ruiz-Trillo I. 2012. Phylogenetic relationships within the Opisthokonta based on phylogenomic analyses of conserved single-copy protein domains. *Mol Biol Evol* **29**:531–544. doi:10.1093/molbev/msr185

Tripathi A, Prem Kumar K V., Chaube SK. 2010. Meiotic cell cycle arrest in mammalian oocytes. *J Cell Physiol*. doi:10.1002/jcp.22108

Ullah Z, Lee CY, Lilly MA, DePamphilis ML. 2009. Developmentally programmed endoreduplication in animals. *Cell Cycle*. doi:10.4161/cc.8.10.8325

Vinayak S, Pawlowic MC, Sateriale A, Brooks CF, Studstill CJ, Bar-Peled Y, Cipriano MJ, Striepen B. 2015. Genetic modification of the diarrhoeal pathogen Cryptosporidium parvum. *Nature* **523**:477–480. doi:10.1038/nature14651

Visintin R, Prinz S, Amon A. 1997. CDC20 and CDH1: A family of substrate-specific activators of APC- dependent proteolysis. *Science* **278**:460–463.

doi:10.1126/science.278.5337.460

Wade Harper J, Adami GR, Wei N, Keyomarsi K, Elledge SJ, Harper JW, Adami GR, Wei N, Keyomarsi K, Elledge SJ. 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **75**:805–816. doi:10.1016/0092-8674(93)90499-G

Wang DZ. 2008. Neurotoxins from marine dinoflagellates: A brief review. *Mar Drugs*. doi:10.3390/md6020349

Wang DZ, Zhang YJ, Zhang SF, Lin L, Hong HS. 2013. Quantitative Proteomic Analysis of Cell Cycle of the Dinoflagellate Prorocentrum donghaiense (Dinophyceae). *PLoS One* **8**. doi:10.1371/journal.pone.0063659

Wang H, Carey LB, Cai Y, Wijnen H, Futcher B. 2009. Recruitment of Cln3 cyclin to promoters controls cell cycle entry via histone deacetylase and other targets. *PLoS Biol*. doi:10.1371/journal.pbio.1000189

Weigmann K, Lehner CF. 1995. Cell fate specification by even-skipped expression in the Drosophila nervous system is coupled to cell cycle progression. *Development* **121**:3713–3721.

Weijer CJCJ, Duschl G, David CN. 1984. A revision of the Dictyostelium discoideum cell cycle. *J Cell Sci* **70**:111–131.

Wheeler RJ, Scheumann N, Wickstead B, Gull K, Vaughan S. 2013. Cytokinesis in trypanosoma brucei differs between bloodstream and tsetse trypomastigote forms: Implications for microtubule-based morphogenesis and mutant analysis. *Mol Microbiol*. doi:10.1111/mmi.12436

White MW, Suvorova ES. 2018. Apicomplexa Cell Cycles: Something Old, Borrowed, Lost, and New. *Trends Parasitol* **34**:759–771. doi:10.1016/j.pt.2018.07.006

Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE, Matese JC, Perou CM, Hurt MM, Brown PO, Botstein D. 2002. Identification of Genes Periodically Expressed in the Human Cell Cycle and Their Expression in Tumors. *Mol Biol Cell* **13**:1977–2000. doi:10.1091/mbc.02-02-0030.

Wilsker D, Petermann E, Helleday T, Bunz F. 2008. Essential function of Chk1 can be

uncoupled from DNA damage checkpoint and replication control. *Proc Natl Acad Sci* **105**:20752–20757. doi:10.1073/pnas.0806917106

Woo YH, Ansari H, Otto TD, Linger CMK, Olisko MK, Michálek J, Saxena A, Shanmugam D, Tayyrov A, Veluchamy A, Ali S, Bernal A, Del Campo J, Cihlář J, Flegontov P, Gornik SG, Hajdušková E, Horák A, Janouškovec J, Katris NJ, Mast FD, Miranda-Saavedra D, Mourier T, Naeem R, Nair M, Panigrahi AK, Rawlings ND, Padron-Regalado E, Ramaprasad A, Samad N, Tomčala A, Wilkes J, Neafsey DE, Doerig C, Bowler C, Keeling PJ, Roos DS, Dacks JB, Templeton TJ, Waller RF, Lukeš J, Oborník M, Pain A. 2015. Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. *Elife* **4**:1–41. doi:10.7554/eLife.06974

Worley AC, Raper KB, Hohl M. 1979. Fonticula alba: A New Cellular Slime Mold (Acrasiomycetes). *Mycologia*. doi:10.2307/3759186

Woznica A, Cantley AM, Beemelmanns C, Freinkman E, Clardy J, King N. 2016. Bacterial lipids activate, synergize, and inhibit a developmental switch in choanoflagellates. *Proc Natl Acad Sci U S A* **113**:7894–7899. doi:10.1073/pnas.1605015113

Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet* **8**:206–216. doi:10.1038/nrg2063

Wu Z, Liu J, Yang H, Xiang H. 2014. DNA replication origins in archaea. *Front Microbiol* **5**:1–7. doi:10.3389/fmicb.2014.00179

Xiong Y, Zhang H, Beach D. 1992. D type cyclins associate with multiple protein kinases and the DNA replication and repair factor PCNA. *Cell* **71**:505–514. doi:10.1016/0092-8674(92)90518-H

Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. 2014. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One* **9**. doi:10.1371/journal.pone.0078644

Zhu W, Giangrande PH, Nevins JR. 2004. E2Fs link the control of G1/S and G2/M transcription. *EMBO J*. doi:10.1038/sj.emboj.7600459

Zhuang Y, Zhang H, Lin S. 2013. Cyclin B gene and its cell cycle-dependent differential

expression in the toxic dinoflagellate Alexandrium fundyense Atama Group I/Clade I. *Harmful Algae* **26**:71–79. doi:10.1016/j.hal.2013.04.002

Zones JM, Blaby IK, Merchant SS, Umen JG. 2015. High-Resolution Profiling of a Synchronized Diurnal Transcriptome from Chlamydomonas reinhardtii Reveals Continuous Cell and Metabolic Differentiation. *Plant Cell* **27**:2743–2769. doi:10.1105/tpc.15.00498

Appendix 1: A detailed description of the cell cycle transcriptional program of *Capsaspora owczarzaki* highlights similarities and differences with other opisthokonts

Alberto Perez-Posada [1]

[1]: Institute of Evolutionary Biology. Pg Marítim de la Barceloneta, 37-49, 08003 Barcelona

One of the main regulatory layers of the cell cycle consists of periodic transcriptional regulation. Independent studies have shown that cell cycle periodic transcription is barely conserved across eukaryotes, only comparable at closer evolutionary distances. In Opisthokonta, the group including animals and fungi, much is known about the cell cycle of few model species but many species in between have been poorly characterized. Comparative genomes of unicellular relatives of animals and fungi have shown an unprecedented degree of conservation of key cell cycle regulators for which we still lack a lot of functional information. To investigate the putative cell cycle regulation of *Capsaspora*, a unicellular relative of animals, we performed a curated, semi-automated exploration of the identity of genes found to be periodically expressed. We identified nearly fifty *Capsaspora* orthologs of genes with known roles in the cell cycle of other eukaryotes or exclusively in animal cells, including transcription factors that belong to cell cycle transcriptional networks conserved throughout all eukaryotes. Further comparisons with datasets from other species, while limited, reassure the position of *Capsaspora* as the opisthokont with the most ancestral transcriptional regulation of the cell cycle.

INTRODUCTION

Periodic transcription during the cell cycle is exhibited in many types of eukaryotic cells, from unicellular to multicellular species (Morgan, 2007; Oliva et al., 2005; Spellman et al., 1998; Whitfield et al., 2002). Despite its omnipresence, periodic transcriptional programs are deployed very differently across distantly related eukaryotic species due to their evolutionary constraints (different lifestyles, adaptations, and all the evolutionary events leading to the biology of a species). Similarities are only appreciable at closer evolutionary distances, such as genus or subgroup (Jensen et al., 2006; Kelliher et al., 2016). Interrogating periodic regulation during the cell cycle at different key evolutionary positions, such as the origin of Opisthokonta or animals, might prove useful to investigate the role of cell cycle in different evolutionary transitions.

In section 3.1 of the present thesis, we present a novel protocol for culture synchronization in *Capsaspora* adherent cells, which was used to further characterize the progression of the cell cycle by investigating gene expression dynamics using RNA-seq (Perez-Posada et al., 2019). We recently presented a study of the periodic transcriptional program of *Capsaspora*, corroborating previous observations in former studies of its genome (Perez-Posada et al., 2019; Suga et al., 2013). Specifically, we illustrated the resemblance of cell cycle regulation of *Capsaspora* and animals by showing a conserved temporal order of expression of *Capsaspora* and human cyclins E, A, and B, and the periodic regulation of a dozen of cell cycle regulators. However, the *Capsaspora* genome contains numerous genes that are known to be involved in the regulation of the eukaryotic cell cycle. These include different transcription factors of the E2F, Forkhead and Myb classes, some of which have been lost in other organisms extensively studied such as yeasts (Cross et al., 2011; Medina et al., 2016).

To expand our vision of how *Capsaspora* might regulate progression through the cell cycle, we performed a curated, semi-automated exploration of the periodic transcriptional program of *Capsaspora*. Using different levels of stringency in our analyses of periodicity, we identified two subdivisions of the periodic gene clusters and found nearly fifty genes with known functions in the cell cycle of other eukaryotes or exclusively in animal cells, including numerous transcription factors that belong to cell cycle transcriptional networks conserved throughout all eukaryotes. We expanded our previous comparative analysis to more datasets from more species, which illustrated larger conservation of cell cycle regulation between human cells, budding yeast, and *Capsaspora*. Interestingly, we also found larger similarities with a subset of the periodic transcriptional program of *Dictyostelium discoideum* (Strasser et al., 2012). We interpret these results as *Capsaspora* retaining a more ancestral periodic transcriptional program, while stressing the need of more data from more taxa to infer more accurate evolutionary

relationships in the evolution and conservation of cell cycle regulation.

RESULTS

A high throughput analysis to detect periodicity patterns, using a combined approach of wave-oriented (JTK_CYCLE) and slope-oriented (RAIN) algorithms, allowed the identification of many genes periodically expressed during the cell cycle, including components of the replication and cell division machineries (Perez-Posada et al., 2019). Since there is a continuum of gene expression dynamics during the cell cycle, and because these algorithms assign a score of probability, no clear distinction can be made between "periodic" and "non-periodic" genes. Therefore, this approach usually relies on a cutoff assigned arbitrarily (see (Kelliher et al., 2016) for an example of a study under the same constrictions). As a consequence, a considerable number of transcripts with an oscillatory pattern -or a pattern of differential expression- are not recovered as periodic by our pipeline, and a substantial number of genes are reported as periodic but, upon visual inspection, only show small fluctuations in their otherwise stable expression pattern. While type II errors (false positives) are arguably controlled by multiple-hypothesis test corrections implemented in each of the two algorithms (Hughes et al., 2010; Thaben and Westermark, 2014), it does not prove enough to recover many genes that could be argued to be periodically expressed. For example, in our analysis, we failed to recover cyclin A gene (CAOG_04719), which clearly coexpresses with genes in the S cluster, around the time where *Capsaspora* cells are replicating their DNA.

To overcome the limitations of these algorithms upon the detection of false and true positives, two additional we followed two additional approaches to identify as many periodic genes as possible. First, we considered a second criterion of "high expression variability", measured as the fold change (FC) in gene expression of a gene throughout the cell cycle. Genes with a fold change in gene expression larger than 1.4 were selected to follow the pipeline of periodicity detection, also applying a cutoff to the top 800. Overall, 50% of the genes were detected as periodic by both the FC-filtered and the original dataset (Fig. A1.1A). The dataset of highly dynamically expressed genes contains a larger number of S phase genes compared to the original, and more genes that cluster in the G1/S cluster (Fig. A1.1B). Further analysis revealed the G1/S cluster to be subdivided in two; one where there is no expression after mitosis, and another where expression rises again after cell division. The latter is nonetheless enriched in genes related to DNA replication as well (Fig. A1.1C). Overall, similar results are obtained in the rest of the analyses. Second, we performed a manual, blindfold revision of gene expression profiles to select transcripts poorly ranked but otherwise clearly differentially regulated genes, from rank 801 in the original dataset down to rank ~2000.

We combined the sets of periodic genes detected by each method while keeping out the 'flat genes' with low fold change (Fig. A1.2A). The resulting list of ~1100 genes constitutes a more permissive dataset that was analysed following the same methodology (Fig. A1.2B). Results from reanalysis using this dataset yielded overall similar results, with similar PCA/PCoA structure and variance (Fig. A1.2C). Six gene clusters were retrieved, with two clearly separated clusters for G2/early mitosis and late mitosis (Fig. A1.2D). This subcluster allows to observe a temporal separation between expression of genes related to chromosome condensation and spindle-kinetochore interaction (Fig. A1.2E). We previously reported an enrichment of tRNA synthesis genes during G2. Interestingly, the G2/early M cluster is significantly enriched in amino acid biosynthesis, another process related to protein synthesis (Fig. A1.2E). t-SNE plotting of this dataset shows a more interconnected layout, compared to the one reported in our manuscript, suggesting more genes with different peaks of expression throughout the cell cycle are retrieved (Fig. A2.2F). Due to this continuity, distinction between cluster borders becomes a bit fuzzier than previously reported. Global pan-eukaryotic-age enrichment decreases slightly, as shown in Fig. A1.2G, meaning more genes innovated at later stages in the evolutionary history of *Capsaspora* are retrieved.

Identification of genes in the *Capsaspora* periodic transcriptional program

While a high-throughput analysis proves useful for a broad overview, we checked the identity of the genes periodically -or differentially- expressed throughout the cell cycle for a more detailed depiction of the cell cycle regulation in *Capsaspora*. From the dataset of 1100 genes, we took those that match a *H. sapiens* ortholog -or have a *H. sapiens* co-ortholog- with a known role in the cell cycle, by any of the three following methods: OrthoFinder assignment (Emms and Kelly, 2015), BLAST top hit alignment (Altschul et al., 1990), or one-to-one ortholog relationships as in (Sebé-Pedrós et al., 2016). A quick browse of BLAST alignments against the *S. cerevisiae* proteome showed that many of these *Capsaspora* genes are also conserved in yeasts (data not shown). Some of these genes were not reported as periodic in our original manuscript, but belong to the dataset later curated.

These genes can be classified in different categories according to their roles and involvement in different aspects of the cell cycle. Expression pattern of these genes can be found at figure 3; for simplicity, we will refer to them as their human orthologs, or co-orthologs.

Entry into cell cycle and G1/S regulators

Several of these genes have human co-orthologs with roles in promoting cell proliferation, such as AGTRAP (CAOG_07352) (Kim et al., 2015), Syntenin (Kashyap et al., 2015), and one ortholog from the MAGE gene family (CAOG_03494) (Barker and Salehi, 2002; Doyle et al., 2010). There is also peak expression of different protein-kinase domain-containing genes belonging to the same orthogroups as MAP kinases, such as MEK (CAOG_05490). We were not able to detect any periodic gene related to CDKN2a, the CDK inhibitor that regulates G1 cyclin-CDK complexes (Bockstaele et al., 2006; Okamoto et al., 1994).

Activation of the cell cycle program is triggered by the de-inhibition of the E2F1 transcription factor by Rb proteins. Whereas periodic activity of E2F1 has not been detected in *Capsaspora*, Rb expression (CAOG_07372) is

downregulated upon entry in S phase, matching the activation of E2F1. Another protein that regulates E2F1 is E2F7, an atypical E2F factor which represses transcription of E2F1 (Aksoy et al., 2012; Di Stefano et al., 2003; Endo-Munoz et al., 2009). Interestingly, *Capsaspora* E2F7 (CAOG_00276) levels also decrease upon entry into S phase. E2F1 cooperates with the transcription factor DP1 to transcribe cell cycle genes. No *Capsaspora* ortholog of DP1 has been found to be periodically expressed.

DNA replication

During S phase, many genes related to DNA replication are upregulated. Different DNA polymerase subunits, including POLA1, POLA2, POLD, and POLE2E are upregulated, together with helicases such as CDC45 (CAOG_00155) (Saha et al., 1998), DNA primases (CAOG_07825), and other proteins that form pre-replicative complexes. Loading of PCNA and other pre-replication proteins may be helped by DCC1 (CAOG_04241) (Merkle et al., 2003), as well as coupling chromatid cohesion and DNA (Terret et al., 2009). Peaks of expression of these genes correlates with upregulation of *Capsaspora* cyclins E and A and the TIMELESS/TIPIN complex (CAOG 00958/04913), which are known to start and stabilize DNA replication by regulating the replisome (Yoshizawa-Sugata and Masai, 2007). Concomitant to this, chromatin proteins are transcribed and synthesized during S phase, such as histones, or the cohesin complex SMC1B (CAOG_04858) and SMC3 (CAOG_01984), members of the cohesin complex keeping sister chromatids together (Sumara et al., 2000).

During this time, damages in the DNA are thought to occur because of hydroxyurea stress (Fox et al., 1987; Singh et al., 2017) and because of stalled replication forks upon coordination with transcription and other processes (Branzei and Foiani, 2005). Different proteins related to DNA damage, such as recombinases Lim15 (CAOG_07267) and Spo11 (CAOG_03056) (Robert et al., 2016), RAD51 (CAOG_01607) (Park et al., 2008; Sigurdsson et al., 2002), MSH2 (CAOG_03813), or PARPs (CAOG_03714) (Couto et al., 2011; Pears and Lakin, 2014; Satoh and Lindahl, 1992), are upregulated in *Capsaspora*. Interestingly, the telomerase complex component p240 (CAOG_01731) peaks at G2, suggesting chromosome termini finish replicating at G2 (Venteicher et al., 2009).

G2 phase genes

The G2 phase serves as a gap phase where the cell monitors a proper completion of former events, such as the end of replication or the synthesis of more proteins to participate in cell division. We detected an ortholog of the Cyclin G/I family in *Capsaspora* (CAOG_00179). Cyclin G is recruited to promote p53-mediated cell cycle arrest in G2/M in response to DNA damage and stress (S. H. Kimura et al., 2001), whereas cyclin I does not oscillate transcriptionally in animals but has been shown to mediate the G2/M transition (Nagano et al., 2013). The oscillatory pattern in G2 suggests that Capsaspora cyclin G/I might operate by delaying the entry into mitosis, similar to the role found in animals. An ubiquitin ligase from the cullin family, CUL3 (CAOG_04201), peaks at G2 putatively regulating protein turnover from the previous cell cycle phases, as happens to cyclin E or Aurora B in human cells (Singer et al., 1999; Sumara et al., 2007). Another putative broader regulator of the cell cycle could be in the *Capsaspora* ortholog of GTP-binding protein 1 (CAOG_00188), known to regulate mRNA stability in a circadian manner (Woo et al., 2011).

We found coexpression of some genes indirectly related to protein synthesis. DDX1 (CAOG_02038) is an RNA helicase reported in many different processes, including activation of transcription factors and cyclins. Together with archease (CAOG_07121), another gene with a peak of expression in G2, they facilitate the guanylation of RTCB (CAOG_03064), a key intermediate step in tRNA ligation which is also upregulated in G2 (Popow et al., 2014). This co-regulation of three genes in the same pathway could be related to the enrichment in GO terms for tRNA maturation during G2 in Capsaspora. tRNA maturation is related to protein translation, which could occur at large amounts during G2/M likely to synthesize the numerous cytoskeletal components required for mitosis.

Cytoskeletal dynamics

Several cytoskeletal proteins involved in actin dynamics in eukaryotes, have a periodically expressed co-ortholog in *Capsaspora*, such as fascin (CAOG_00505) (Ma et al., 2013), filopodial upregulator HBV (CAOG_03349) (Dasgupta et al., 2011), drebrin (CAOG_08639) (Pérez-Martínez et al., 2010), and thymosin beta 4 (CAOG_06989) (SCHÖBITZ et al., 1991). The latter, known to bind and sequester g-actin preventing polymerization of actin filaments in human cells, is by far the most highly expressed gene in the *Capsaspora* cell cycle accounting to up to 15% of the total transcript pool at its peak of expression, at times corresponding to G2/M stages. Being this the moment where the cells round up their shape to divide, it is tempting to speculate on TYB4 as one of the main drivers of changes in the cytoskeleton of *Capsaspora*. CLASP1 (CAOG_03237), a protein previously reported to stabilize dynamic microtubules (Akhmanova et al., 2001; Efimov et al., 2007; Mimori-Kiyosue et al., 2005), peaks at G2/M, putatively contributing to stabilization of spindle microtubules. We detect a dynamic behavior of integrins A1 (CAOG_01284), B1 (CAOG_01283), and Vinculin (CAOG_05123), with a pre-replicative differential expression pattern. As the integrin adhesome is known to be involved in cell-extracellular matrix adhesion in human cells (Ruoslahti and Giancotti, 1989) and putatively in *Capsaspora* (Parra-Acero et al., in prep.), this might indicate that cell adhesion decreases upon entry into mitosis, which matches our observations of cells retracting filopodia, dettaching, and rounding up prior to dividing.

Spindle and chromosome interaction, chromosome separation, and mitosis

During cell division, different mechanisms and checkpoints ensure that chromosomes are properly aligned and positioned at midcell prior to their segregation. Sorting nexin-33 (CAOG_04878), suggested to have a role in chromosome alignment and segregation (Ma and Chircop, 2012), is periodically expressed during G2 in Capsaspora. We found mitotic expression of HASPIN (CAOG_09033), a protein kinase associated to histone 3 which mediates the positioning of AURKB during mitosis (Dai, 2005). Condensin subunits SMC2(00350) and SMC4

(CAOG_01474), which condense the newly replicated sister chromatids (K. Kimura et al., 2001), also peak in M phase in *Capsaspora*. Together with these genes, we found numerous co-orhtologs of kinesins, dineins, centromeric proteins (reported in the manuscript), and DNA damage proteins such as RAD54 (CAOG_04682) (Ristic et al., 2001; Sigurdsson et al., 2002; Swagemakers et al., 1998), peaking at mitosis.

Overall, *Capsaspora* seems to have a fairly conserved set of proteins regulating the mitotic cell cycle, with representative members of each major checkpoint complex (entry, G2/M, spindle, and mitotic checkpoints), that are cell cycle-regulated at the transcriptional level. We identified periodic co-orthologues of proteins Rod (CAOG_03064) and Zwilch (CAOG_07213), which form the RZZ complex together with the gene ZW10, involved in the mitotic spindle checkpoint (Kops et al., 2005; Samerier, 2011). Although not detected as periodic, there also exists a co-orthologue of ZW10 in *Capsaspora* (CAOG_00245). The RZZ complex interacts with the protein Spindly, reported to recruit dynein into kinetochores and proposed to contribute to the regulation of the spindle checkpoint (Griffis et al., 2007). A co-orthologue of Spindly in *Capsaspora* (CAOG_02305) also presents periodic transcription, peaking at division times together with Rod and Zwilch orthologues. We also identified proteins from the NDC80 complex, such as NDC80 (CAOG_00634) and SP24 (CAOG_003380), required to the organization of spindle tips and kinetochore complexes (McCleland et al., 2004), and an ortholog of BUB1 kinase, essential for localization and assembly of checkpoint proteins at the kinetochore (Johnson et al., 2004; Klebig et al., 2009; Qi and Yu, 2007; Seeley et al., 1999; Tang et al., 2004), is also expressed at later stages of the *Capsaspora* cell cycle. Cdc25 is one of the main regulators of the cell cycle, as dephosphorylates mitotic cyclin-CDK complexes and activates them to progress through M phase (Peng et al., 1997; Sadhu et al., 1990). The single *Capsaspora* ortholog of the same group as human CDC25 is differentially regulated throughout the cell cycle (CAOG_03819), with the highest peaks at the of the cell cycle. A *Capsaspora* Tribbles homolog (CAOG_01948), a regulator in charge of the protein turnover of cdc25 (Liang et al., 2016), has a clear peak of expression in G2/M. We hypothesize that these genes might function in the spindle and mitotic checkpoints of *Capsaspora*.

Other *Capsaspora* regulators peaking at mitosis are GSTP1 (CAOG_02067), the absence of which delays entry and progression through G1 and postpones the expression of G1 cell cycle regulators in human cells (Pajaud et al., 2015), and CKS (CAOG_07008), a modulator of cyclin-CDK activity. At the onset of G1, organelle dynamics may increase provided the G1 upregulation of TWNK (CAOG_09482) and mitofusin (CAOG_00950), regulating replication of mtDNA and mitochondrion dynamics (Gupte, 2015; Korhonen et al., 2004).

Transcription factors during the cell cycle of *Capsaspora*

To assess the putative identity of the transcription factors (TFs) controlling the periodic transcriptional program of *Capsaspora*, we proceeded to characterize the putative co-expression of transcription factors throughout the cell cycle. To that end, we checked which *Capsaspora* genes containing DNA binding domains were consistently expressed along with the two experimental replicates. From a total of 168 genes with DNA-binding domains, we inferred groups of co-expression by Pearson correlation and hierarchical clustering, followed up by manual testing (Fig. A1.4A) Two main groups of TF co-expression were detected: the G1/S transition, G2, and G1 expression. At G1/S, there are transcription factors such as SOX (CAOG_04975T0), SIX1 (CAOG_05435) and CSRP3(CAOG_04350), related to myogenesis and cell proliferation in animals, the former known to upregulate cyclin A expression. MAX (CAOG_01784) expression decreases two-fold during S and G2, and correlates to NFY (CAOG_04030), CG1, TFIIA, and GATA domain-containing proteins (CAOG_06632, CAOG_00755, CAOG_00358). Proteins from the GATA family (CAOG_03836) are also expressed in G2, together with nuclear factor X (CAOG_06498). These clusters anti-correlate with a small set of genes with low expression variability throughout the cell cycle, although visual inspection suggests they might be differentially expressed at later stages. Among these, the mitotic transcription factor FoxM1 (CAOG_05750) is found.

In addition, DNA motif enrichment analysis was performed in the promoters of periodic genes within each independent cluster compared to the rest of promoters using HOMER Software (Heinz et al., 2010). These analyses yielded putative DNA motifs for each cluster showing two- to five-fold enrichment (Fig. A1.4B, Table 1). Among these, one motif in promoters of the G1/S cluster showed similarity to the motif found for animal DP1 transcription factor, which couples with E2F during the entry into cell cycle (Helin et al., 1993; Wu et al., 1995). No enriched motif was found that matches the E2F consensus binding site; nevertheless, it has been estimated that only a small fraction of E2F-mediated transcription occurs by direct binding to DNA (Rabinovich et al., 2008); the rest occurring indirectly mediated by other transcription factors such as DP. Similarly, the M cluster shows enrichment of a DNA motif resembling B-MYB binding site, another transcription factor known to upregulate genes at mitosis (Fischer and Müller, 2017). One MYB-related protein from *Capsaspora* (CAOG_00301) peaks at S phase, preceding its putative control of mitotic genes (Fig. A1.4B). Another putative motif resembling the NFX binding site, a TF differentially expressed during the cell cycle, was also found enriched in the G2/M cluster. A full list of motifs can be found in Table 1.

All together, these findings suggest that the mitotic gene cluster might be controlled by the activity of Myb-like transcription factors, as previously described for other species, and that upregulation of G1/S genes driven cooperatively by different TFs. Despite these results, more functional experiments of protein-DNA interaction would be required for further insight.

Comments on the *Capsaspora* cyclin-CDK content in the context of Holozoa and other eukaryotes

Previous studies in the literature showed inconsistencies regarding the presence of a CDK1 ortholog in *Capsaspora* (Cao et al., 2014; Suga et al., 2013). To unravel the reason

behind, we decided to characterize the repertoire of cyclins and CDKs in *Capsaspora* using phylogenetic trees. We surveyed the proteomes of a large set of species representing all the main eukaryotic lineages.

Whereas other studies claim to have found A,B,D,E-type cyclins in all the major eukaryotic lineages (Medina et al., 2016), we were not able to determine the phylogenetic relationships of these subfamilies in other groups outside Opisthokonta due to the poor phylogenetic signal in this protein family. We did, however, detected an ortholog of CDK1 in *Capsaspora* (CAOG_07905), previously unreported in other studies (Fig. A1.5, see Result 3.1). This could be explained if an e-value constraint was taken into account when identifying CDK proteins in (Cao et al., 2014), as the gene CAOG_07905 was misannotated due to a gap in the genome assembly. We were able to identify and validate the rest of the coding sequence by PCR from gDNA and cDNA, as well as our transcriptomic data. The full sequence of CAOG_07905 aligns to the complete sequence of human and yeast CDK1 and cdc28, respectively (see Result 3.2).

Our analysis confirms previous findings about the expansion of different cyclin subfamilies before Holozoa in a deeper branch of the tree of eukaryotes with the exception of cell cycle cyclins D and E. In the former literature, Archaeplastid "D-type" cyclins form a monophyletic group with metazoan cyclins G/I; this group appears as sister group to metazoan D-type cyclins (Medina et al., 2016). Whereas *Capsaspora* possesses a Cyclin G/I gene, this gene shows a G2 peak (discussed in Result 3.1) and does not follow an expression pattern typical of genes in the G1 clusters. In addition, the single gene from *Corallochytrium limacisporum* retrieved as sister group of metazoan D-type cyclins (Fig. A1.5) might be an artifact of long-branch attraction provided the absence of this cyclin subfamily in the rest of unicellular Holozoa. Similarly, it was reported a group of E-type cyclins in chytrids which seems polyphyletic to animal E-type cyclins, but no other eukaryotic groups are reported to have any member of these families.

Comparison with other datasets

Our comparisons with the periodic transcriptional program (PTP) of other species shows that the *Capsaspora* PTP is more similar with human cells PTP. As discussed in the manuscript of Section 3.1, comparative approaches are difficult for numerous reasons: aiming to reduce the cumulative bias due to distantly related species, differences in detection methods, and , we sought to further investigate our findings in order to clarify our observations. We did so by extending our comparative analysis to other datasets of periodic genes from the literature, including more human and yeast datasets (Oliva et al., 2005; Peña-Diaz et al., 2013; Pramila et al., 2006; Spellman et al., 1998), a dataset of Cryptococcus neoformans (Kelliher et al., 2016), and a data set of potentially G1/S-, G2-, and M-regulated genes in *D. discoideum* (Strasser et al., 2012).

We found an expanded trend of our previous observations, with small but nonetheless significant numbers of periodic orthogroups shared between species. Adding more yeast datasets confirms the tendency of conserved genes between programs from the same species, as illustrated by the highlighted similarities between human and yeasts (Fig. A1.6). We can see a larger-than-expected conservation of a small number of periodic genes between *D. discoideum*, *Capsaspora*, and yeasts. These similarities are interpreted as indeed a conservation of the periodic transcriptional program because comparisons using randomized gene lists for each species yielded very different, non-significant results (see Fig A1.7A for an example), thus discarding a byproduct of phylogenetic closeness between species (or datasets of the same species) and pointing to the presence of signal in the data.

With the availability of more additional datasets, we also did additional comparisons of *Capsaspora* with other sets of core cell cycle genes, including one of yeasts (resulting from the intersection of three different budding yeast datasets). Interestingly, we detected that *Capsaspora* shares a remarkably large amount of periodic orthogroups with *S. cerevisiae*, the same happening between human cells and *S. cerevisiae* (Fig A1.7B).

Discussion

The discussion of this Appendix is included in the Section 4.1 of the present thesis. Briefly, these results prompt to suggest that the transcriptional regulation of the *Capsaspora* cell cycle is largely similar to that of other Opisthokonta, perhaps retaining a more ancestral state provided the similarities with an outgroup species such as *D. discoideum*. One possible interpretation of our comparative analyses is that *Capsaspora*, *S. cerevisiae*, and *H. sapiens* all share a more conserved periodic transcriptional program than the rest of species in this study, despite the otherwise divergent regulatory mechanisms during the cell cycle. Our lack of data from other fungi outside unicellular Ascomycota and from other animals hampers to consider conservation or convergence in the use of periodic gene expression as a mean to regulate progression through the cell cycle between these species. All together, these results suggest that pairwise species comparison alone performs poorly as a proxy of how similar is cell cycle regulation between two species compared to a third species, requiring more data at the post-translational level to infer the similarities in protein activity during the cell cycle to complement our current data at the transcriptional level (see Section 4.1 of the present thesis).

References

Akhmanova A, Hoogenraad CC, Drabek K, Stepanova T, Dortland B, Verkerk T, Vermeulen W, Burgering BM, De Zeeuw CI, Grosveld F, Galjart N. 2001. CLASPs Are CLIP-115 and -170 Associating Proteins Involved in the Regional Regulation of Microtubule Dynamics in Motile Fibroblasts. *Cell* **104**:923–935. doi:10.1016/S0092-8674(01)00288-4

Aksoy O, Chicas A, Zeng T, Zhao Z, McCurrach M, Wang X, Lowe SW. 2012. The atypical E2F family member E2F7 couples the p53 and RB pathways during cellular senescence. *Genes Dev* **26**:1546–1557. doi:10.1101/gad.196238.112

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol*.

doi:10.1016/S0022-2836(05)80360-2

Barker PA, Salehi A. 2002. The MAGE proteins: Emerging roles in cell cycle progression, apoptosis, and neurogenetic disease. *J Neurosci Res* **67**:705–712. doi:10.1002/jnr.10160

Bockstaele L, Kooken H, Libert F, Paternot S, Dumont JE, de Launoit Y, Roger PP, Coulonval K. 2006. Regulated Activating Thr172 Phosphorylation of Cyclin-Dependent Kinase 4(CDK4): Its Relationship with Cyclins and CDK "Inhibitors." *Mol Cell Biol* **26**:5070–5085. doi:10.1128/mcb.02006-05

Branzei D, Foiani M. 2005. The DNA damage response during DNA replication. *Curr Opin Cell Biol* **17**:568–575. doi:10.1016/j.ceb.2005.09.003

Cao L, Chen F, Yang X, Xu W, Xie J, Yu L. 2014. Phylogenetic analysis of CDK and cyclin proteins in premetazoan lineages. *BMC Evol Biol*. doi:10.1186/1471-2148-14-10

Couto CA-M, Wang H-Y, Green JCA, Kiely R, Siddaway R, Borer C, Pears CJ, Lakin ND. 2011. PARP regulates nonhomologous end joining through retention of Ku at double-strand breaks. *J Cell Biol* **194**:367–375. doi:10.1083/jcb.201012132

Cross FR, Buchler NE, Skotheim JM. 2011. Evolution of networks and sequences in eukaryotic cell cycle control. *Philos Trans R Soc B Biol Sci* **366**:3532–3544. doi:10.1098/rstb.2011.0078

Dai J. 2005. The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. *Genes Dev* **19**:472–488. doi:10.1101/gad.1267105

Dasgupta S, Cushman I, Kpetemey M, Casey PJ, Vishwanatha JK. 2011. Prenylated C17orf37 induces filopodia formation to promote cell migration and metastasis. *J Biol Chem* **286**:25935–25946. doi:10.1074/jbc.M111.254599

Di Stefano L, Jensen MR, Helin K. 2003. E2F7, a novel E2F featuring DP-independent repression of a subset of E2F-regulated genes. *EMBO J* **22**:6289–6298. doi:10.1093/emboj/cdg613

Doyle JM, Gao J, Wang J, Yang M, Potts PR. 2010. MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. *Mol Cell* **39**:963–974. doi:10.1016/j.molcel.2010.08.029

Efimov A, Kharitonov A, Efimova N, Loncarek J, Miller PM, Andreyeva N, Gleeson P, Galjart N, Maia ARR, McLeod IX, Yates JR, Maiato H, Khodjakov A, Akhmanova A, Kaverina I. 2007. Asymmetric CLASP-Dependent Nucleation of Noncentrosomal Microtubules at the trans-Golgi Network. *Dev Cell* **12**:917–930. doi:10.1016/j.devcel.2007.04.002

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* **16**:157. doi:10.1186/s13059-015-0721-2

Endo-Munoz L, Dahler A, Teakle N, Rickwood D, Hazar-Rethinam M, Abdul-Jabbar I, Sommerville S, Dickinson I, Kaur P, Paquet-Fifield S, Saunders N. 2009. E2F7 can regulate proliferation, differentiation, and apoptotic responses in human keratinocytes: Implications for cutaneous squamous cell carcinoma formation. *Cancer Res* **69**:1800–1808. doi:10.1158/0008-5472.CAN-08-2725

Fischer M, Müller GA. 2017. Cell cycle transcription control: DREAM/MuvB and RB-E2F complexes. *Crit Rev Biochem Mol Biol* **52**:638–662. doi:10.1080/10409238.2017.1360836

Fox MH, Read RA, Bedford JS. 1987. Comparison of synchronized Chinese hamster ovary cells obtained by mitotic shake-off, hydroxyurea, aphidicolin, or methotrexate. *Cytometry*. doi:10.1002/cyto.990080312

Griffis ER, Stuurman N, Vale RD. 2007. Spindly, a novel protein essential for silencing the spindle assembly checkpoint, recruits dynein to the kinetochore. *J Cell Biol* **177**:1005–1015. doi:10.1083/jcb.200702062

Gupte TM. 2015. Mitochondrial fragmentation due to inhibition of fusion increases cyclin B through mitochondrial superoxide radicals. *PLoS One* **10**:1–19. doi:10.1371/journal.pone.0126829

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell* **38**:576–589. doi:10.1016/j.molcel.2010.05.004

Helin K, Wu CL, Fattaey AR, Lees JA, Dynlacht BD, Ngwu C, Harlow E. 1993. Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperative trans-activation. *Genes Dev* **7**:1850–1861. doi:10.1101/gad.7.10.1850

Hughes ME, Hogenesch JB, Kornacker K. 2010. JTK-CYCLE: An efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms* **25**:372–380. doi:10.1177/0748730410379711

Jensen LJ, Jensen TS, De Lichtenberg U, Brunak S, Bork P. 2006. Co-evolution of transcriptional and post-translational cell-cycle regulation. *Nature* **443**:594–597. doi:10.1038/nature05186

Johnson VL, Scott MIF, Holt S V., Hussein D, Taylor SS. 2004. Bub1 is required for kinetochore localization of BubR1, Cenp-E, Cenp-F and Mad2, and chromosome congression. *J Cell Sci* **117**:1577–1589. doi:10.1242/jcs.01006

Kashyap R, Roucourt B, Lembo F, Fares J, Carcavilla AM, Restouin A, Zimmermann P, Ghossoub R. 2015. Syntenin controls migration, growth, proliferation, and cell cycle progression in cancer cells. *Front Pharmacol* **6**.

doi:10.3389/fphar.2015.00241

Kelliher CM, Leman AR, Sierra CS, Haase SB. 2016. Investigating Conservation of the Cell-Cycle-Regulated Transcriptional Program in the Fungal Pathogen, Cryptococcus neoformans. *PLoS Genet* **12**:1–23. doi:10.1371/journal.pgen.1006453

Kim CS, Kim IJ, Bae EH, Ma SK, Lee J, Kim SW. 2015. Angiotensin-(1-7) Attenuates Kidney Injury Due to Obstructive Nephropathy in Rats. *PLoS One* **10**:e0142664. doi:10.1371/journal.pone.0142664

Kimura K, Cuvier O, Hirano T. 2001. Chromosome Condensation by a Human Condensin Complex in Xenopus Egg Extracts. *J Biol Chem* **276**:5417–5420. doi:10.1074/jbc.C000873200

Kimura SH, Ikawa M, Ito A, Okabe M, Nojima H. 2001. Cyclin G1 is involved in G2/M arrest in response to DNA damage and in growth control after damage recovery. *Oncogene* **20**:3290–3300. doi:10.1038/sj.onc.1204270

Klebig C, Korinth D, Meraldi P. 2009. Bub1 regulates chromosome segregation in a kinetochore-independent manner. *J Cell Biol* **185**:841–858. doi:10.1083/jcb.200902128

Kops GJPL, Kim Y, Weaver BAA, Mao Y, McLeod I, Yates JR, Tagaya M, Cleveland DW. 2005. ZW10 links mitotic checkpoint signaling to the structural kinetochore. *J Cell Biol* **169**:49–60. doi:10.1083/jcb.200411118

Korhonen JA, Pham XH, Pellegrini M, Falkenberg M. 2004. Reconstitution of a minimal mtDNA replisome in vitro. *EMBO J* **23**:2423–2429. doi:10.1038/sj.emboj.7600257

Liang KL, Paredes R, Carmody R, Eyers PA, Meyer S, McCarthy T V., Keeshan K. 2016. Human TRIB2 oscillates during the cell cycle and promotes ubiquitination and degradation of CDC25C. *Int J Mol Sci* **17**:1–13. doi:10.3390/ijms17091378

Ma MPC, Chircop M. 2012. SNX9, SNX18 and SNX33 are required for progression: Through and completion of mitosis. *J Cell Sci* **125**:4372–4382. doi:10.1242/jcs.105981

Ma Y, Li A, Faller WJ, Libertini S, Fiorito F, Gillespie DA, Sansom OJ, Yamashiro S, Machesky LM. 2013. Fascin 1 is transiently expressed in mouse melanoblasts during development and promotes migration and proliferation. *Dev* **140**:2203–2211. doi:10.1242/dev.089789

McCleland ML, Kallio MJ, Barrett-Wilt GA, Kestner CA, Shabanowitz J, Hunt DF, Gorbsky GJ, Stukenberg PT. 2004. The Vertebrate Ndc80 Complex Contains Spc24 and Spc25 Homologs, which Are Required to Establish and Maintain Kinetochore-Microtubule Attachment. *Curr Biol* **14**:131–137. doi:10.1016/j.cub.2003.12.058

Medina EM, Turner JJ, Gordân R, Skotheim JM, Buchler NE. 2016. Punctuated evolution and transitional hybrid network in an ancestral cell cycle of fungi. *Elife* **5**. doi:10.7554/eLife.09492

Merkle CJ, Karnitz LM, Henry-Sánchez JT, Chen J. 2003. Cloning and Characterization of hCTF18, hCTF8, and hDCC1. *J Biol Chem* **278**:30051–30056. doi:10.1074/jbc.M211591200

Mimori-Kiyosue Y, Grigoriev I, Lansbergen G, Sasaki H, Matsui C, Severin F, Galjart N, Grosveld F, Vorobjev I, Tsukita S, Akhmanova A. 2005. CLASP1 and CLASP2 bind to EB1 and regulate microtubule plus-end dynamics at the cell cortex. *J Cell Biol* **168**:141–153. doi:10.1083/jcb.200405094

Morgan DO. 2007. Cell Cycle: Principles of Control. *Yale J Biol Med*. doi:10.1093/icb/icm066

Nagano T, Hashimoto T, Nakashima A, Hisanaga S, Kikkawa U, Kamada S. 2013. Cyclin I is involved in the regulation of cell cycle progression. *Cell Cycle*. doi:10.4161/cc.25623

Okamoto A, Demetrickt DJ, Spillare EA, Hagiwara K, Perwez Hussain S, Bennett WP, Forrester K, Gerwin B, Serranot M, Beacht DH, Harris CC. 1994. Mutations and altered expression of p16INK4 in human cancer (p53 protein/tumor-suppressor gene/cyclin Di/retinoblastoma protein). *Proc Nail Acad Sci USA* **91**:11045–11049.

Oliva A, Rosebrock A, Ferrezuelo F, Pyne S, Chen H, Skiena S, Futcher B, Leatherwood J. 2005. The Cell Cycle–Regulated Genes of Schizosaccharomyces pombe. *PLoS Biol* **3**:e225. doi:10.1371/journal.pbio.0030225

Pajaud J, Ribault C, Ben Mosbah I, Rauch C, Henderson C, Bellaud P, Aninat C, Loyer P, Morel F, Corlu A. 2015. Glutathione transferases P1/P2 regulate the timing of signaling pathway activations and cell cycle progression during mouse liver regeneration. *Cell Death Dis* **6**:e1598-13. doi:10.1038/cddis.2014.562

Park JY, Yoo HW, Kim BR, Park R, Choi SY, Kim Y. 2008. Identification of a novel human Rad51 variant that promotes DNA strand exchange. *Nucleic Acids Res* **36**:3226–3234. doi:10.1093/nar/gkn171

Pears CJ, Lakin ND. 2014. Emerging models for DNA repair: Dictyostelium discoideum as a model for nonhomologous end-joining. *DNA Repair (Amst)* **17**:121–131. doi:10.1016/j.dnarep.2014.01.008

Peña-Diaz J, Hegre SA, Anderssen E, Aas PA, Mjelle R, Gilfillan GD, Lyle R, Drabløs F, Krokan HE, Sætrom P. 2013. Transcription profiling during the cell cycle shows that a subset of Polycomb-targeted genes is upregulated during DNA replication. *Nucleic Acids Res* **41**:2846–2856. doi:10.1093/nar/gks1336

Peng CY, Graves PR, Thoma RS, Wu Z, Shaw AS, Piwnica-Worms H. 1997. Mitotic and G2 checkpoint control: Regulation of 14-3-3 protein binding by phosphorylation of Cdc25c on serine-216. *Science* **277**:1501–1505. doi:10.1126/science.277.5331.1501

Pérez-Martínez M, Gordón-Alonso M, Cabrero JR, Barrero-Villar M, Rey M, Mittelbrunn M, Lamana A, Morlino G, Calabia C, Yamazaki H, Shirao T, Vázquez J, González-Amaro R, Veiga E, Sánchez-Madrid F. 2010. F-actin-binding protein drebrin regulates CXCR4 recruitment to the immune synapse. *J Cell Sci* **123**:1160–1170. doi:10.1242/jcs.064238

Perez-Posada A, Dudin O, Ocaña-Pallarès E, Ruiz-Trillo I, Ondracka A. 2019. Gradual evolution of cell cycle regulation by cyclin-dependent kinases during the transition to animal multicellularity. *bioRxiv* 719534. doi:10.1101/719534

Popow J, Jurkin J, Schleiffer A, Martinez J. 2014. Analysis of orthologous groups reveals archease and DDX1 as tRNA splicing factors. *Nature*. doi:10.1038/nature13284

Pramila T, Wu W, Miles S, Noble WS, Breeden LL. 2006. The Forkhead transcription factor Hcm1 regulates chromosome segregation genes and fills the S-phase gap in the transcriptional circuitry of the cell cycle. *Genes Dev*. doi:10.1101/gad.1450606

Qi W, Yu H. 2007. KEN-box-dependent degradation of the Bub1 spindle checkpoint kinase by the anaphase-promoting complex/cyclosome. *J Biol Chem* **282**:3672–3679. doi:10.1074/jbc.M609376200

Rabinovich A, Jin VX, Rabinovich R, Xu X, Farnham PJ. 2008. E2F in vivo binding specificity: Comparison of consensus versus nonconsensus binding sites. *Genome Res* **18**:1763–1777. doi:10.1101/gr.080622.108

Ristic D, Wyman C, Paulusma C, Kanaar R. 2001. The architecture of the human Rad54-DNA complex provides evidence for protein translocation along DNA. *Proc Natl Acad Sci* **98**:8454–8460. doi:10.1073/pnas.151056798

Robert T, Nore A, Brun C, Maffre C, Crimi B, Guichard V, Bourbon H-M, de Massy B. 2016. The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science* **351**:943–949. doi:10.1126/science.aad5309

Ruoslahti E, Giancotti FG. 1989. Integrins and tumor cell dissemination. *Cancer Cells*.

Sadhu K, Reed SI, Richardson H, Russell P. 1990. Human homolog of fission yeast cdc25 mitotic inducer is predominantly expressed in G2. *Proc Natl Acad Sci U S A* **87**:5139–5143. doi:10.1073/pnas.87.13.5139

Saha P, Thome KC, Yamaguchi R, Hou ZH, Weremowicz S, Dutta A. 1998. The human homolog of Saccharomyces cerevisiae CDC45. *J Biol Chem* **273**:18205–18209. doi:10.1074/jbc.273.29.18205

Samerier M. 2011. Functional analyses of microtubule and centrosome-associated proteins in Dictyostelium dicsoideum. University of Postdam.

Satoh MS, Lindahl T. 1992. Role of poly(ADP-ribose) formation in DNA repair. *Nature* **356**:356–358. doi:10.1038/356356a0

SCHÖBITZ B, NETZKER R, HANNAPPEL E, BRAND K. 1991. Cell-cycle-regulated expression of thymosin β4 in thymocytes. *Eur J Biochem* **199**:257–262. doi:10.1111/j.1432-1033.1991.tb16118.x

Sebé-Pedrós A, Peña MI, Capella-Gutiérrez S, Antó M, Gabaldón T, Ruiz-Trillo II, Sabidó E, Gabaldon T, Ruiz-Trillo II, Sabidó E. 2016. High-Throughput Proteomics Reveals the Unicellular Roots of Animal Phosphosignaling and Cell Differentiation. *Dev Cell* **39**:186–197. doi:10.1016/j.devcel.2016.09.019

Seeley TW, Wang L, Zhen JY. 1999. Phosphorylation of human MAD1 by the BUB1 kinase in vitro. *Biochem Biophys Res Commun* **257**:589–595. doi:10.1006/bbrc.1999.0514

Sigurdsson S, Van Komen S, Petukhova G, Sung P. 2002. Homologous DNA pairing by human recombination factors Rad51 and Rad54. *J Biol Chem* **277**:42790–42794. doi:10.1074/jbc.M208004200

Singer JD, Gurian-West M, Clurman B, Roberts JM. 1999. Cullin-3 targets cyclin E for ubiquitination and controls S phase in mammalian cells. *Genes Dev* **13**:2375–2387. doi:10.1101/gad.13.18.2375

Singh A, Agarwal A, Xu Y jie. 2017. Novel Cell-killing mechanisms of hydroxyurea and the implication toward combination therapy for the treatment of fungal infections. *Antimicrob Agents Chemother*. doi:10.1128/AAC.00734-17

Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, Futcher B. 1998. Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast Saccharomyces cerevisiae by Microarray Hybridization. *Mol Biol Cell*. doi:10.1091/mbc.9.12.3273

Strasser K, Bloomfield G, MacWilliams A, Ceccarelli A, MacWilliams H, Tsang A. 2012. A retinoblastoma orthologue is a major regulator of S-phase, mitotic, and developmental gene expression in Dictyostelium. *PLoS One* **7**. doi:10.1371/journal.pone.0039914

Suga H, Chen Z, de Mendoza A, Sebé-Pedrós A, Brown MW, Kramer E, Carr M, Kerner P, Vervoort M, Sánchez-Pons N, Torruella G, Derelle R, Manning G, Lang BF, Russ C, Haas BJ, Roger AJ, Nusbaum C, Ruiz-Trillo I. 2013. The Capsaspora genome reveals a complex unicellular prehistory of animals. *Nat Commun* **4**:1–9. doi:10.1038/ncomms3325

Sumara I, Quadroni M, Frei C, Olma MH, Sumara G, Ricci R, Peter M. 2007. A Cul3-Based E3 Ligase Removes Aurora B from Mitotic Chromosomes, Regulating Mitotic Progression and Completion of Cytokinesis in Human Cells. *Dev Cell* **12**:887–900. doi:10.1016/j.devcel.2007.03.019

Sumara I, Vorlaufer E, Gieffers C, Peters BH, Peters JM. 2000. Characterization of vertebrate cohesin complexes and their regulation in prophase. *J Cell Biol* **151**:749–761. doi:10.1083/jcb.151.4.749

Swagemakers SMA, Essers J, De Wit J, Hoeijmakers JHJ, Kanaar R. 1998. The human Rad54 recombinational DNA repair protein is a double-stranded DNA-dependent ATPase. *J Biol Chem* **273**:28292–28297. doi:10.1074/jbc.273.43.28292

Tang Z, Shu H, Oncel D, Chen S, Yu H. 2004. Phosphorylation of Cdc20 by Bub1 provides a catalytic mechanism for APC/C inhibition by the spindle checkpoint. *Mol Cell* **16**:387–397. doi:10.1016/j.molcel.2004.09.031

Terret M-E, Sherwood R, Rahman S, Qin J, Jallepalli P V. 2009. Cohesin acetylation speeds the replication fork. *Nature* **462**:231–234. doi:10.1038/nature08550

Thaben PF, Westermark PO. 2014. Detecting rhythms in time series with rain. *J Biol Rhythms* **29**:391–400. doi:10.1177/0748730414553029

Venteicher AS, Abreu EB, Meng Z, McCann KE, Terns RM, Veenstra TD, Terns MP, Artandi SE. 2009. A Human Telomerase Holoenzyme Protein Required for Cajal Body Localization and Telomere Synthesis. *Science* **323**:644–648. doi:10.1126/science.1165357

Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE, Matese JC, Perou CM, Hurt MM, Brown PO, Botstein D. 2002. Identification of Genes Periodically Expressed in the Human Cell Cycle and Their Expression in Tumors. *Mol Biol Cell* **13**:1977–2000. doi:10.1091/mbc.02-02-0030.

Woo KC, Kim TD, Lee KH, Kim DY, Kim S, Lee HR, Kang HJ, Chung SJ, Senju S, Nishimura Y, Kim KT. 2011. Modulation of exosome-mediated mRNA turnover by interaction of gtp-binding protein 1 (GTPBP1) with its target mrnas. *FASEB J* **25**:2757–2769. doi:10.1096/fj.10-178715

Wu CL, Zukerberg LR, Ngwu C, Harlow E, Lees JA. 1995. In vivo association of E2F and DP family proteins. *Mol Cell Biol* **15**:2536–2546. doi:10.1128/mcb.15.5.2536

Yoshizawa-Sugata N, Masai H. 2007. Human Tim/Timeless-interacting protein, Tipin, is required for efficient progression of S phase and DNA replication checkpoint. *J Biol Chem*. doi:10.1074/jbc.M605596200

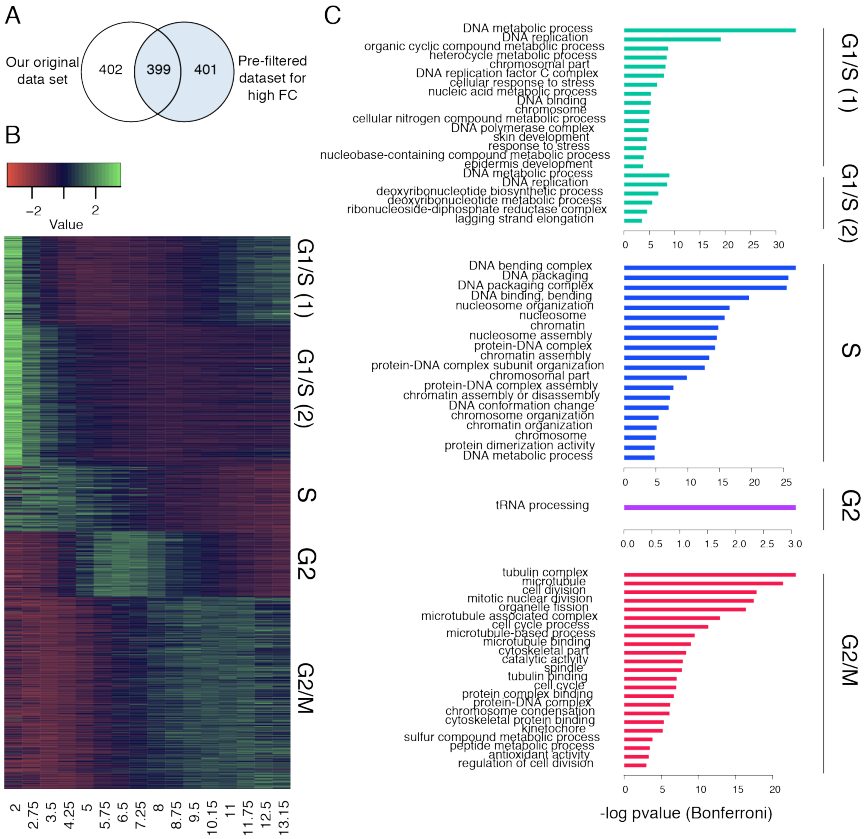


Figure A1.1: Clustering and Gene Ontology enrichment analysis of a set of highly-dynamically expressed *Capsaspora* periodic genes. A: Venn diagram indicating the number of periodic genes detected in the former analysis and in the pre-filtered dataset for high foldchange of expression. B: heatmap showing gene expression across time for the five largest clusters retrieved by hierarchical clustering. C: Barplot showing all the significantly enriched Gene Ontology terms from the five clusters depicted in B.

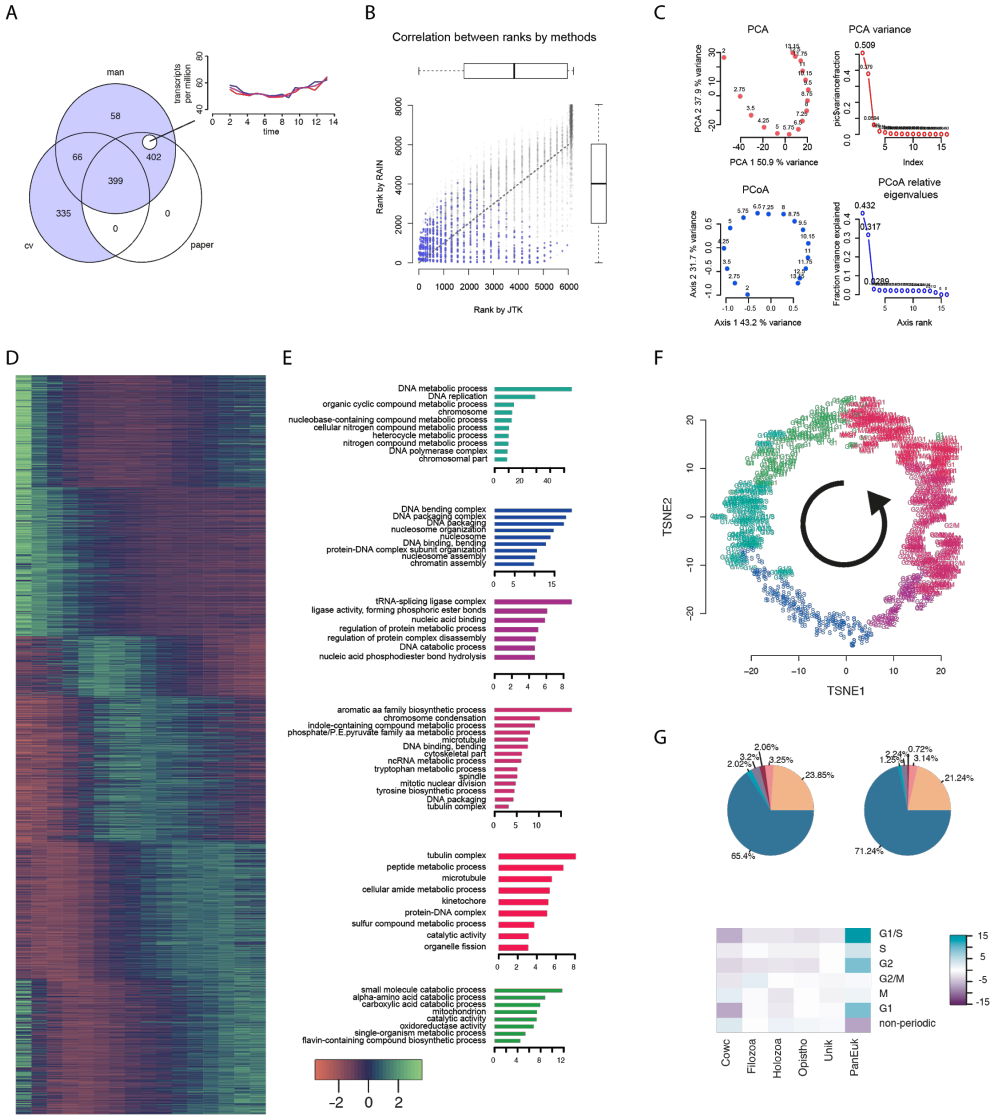


Figure A1.2: Reanalysis of the *Capsaspora* periodic transcriptional program using a more permissive dataset of roughly 1100 genes. A: Venn diagram showing the genes used for the reanalysis. Genes not present in the dataset of high-FC (fold-change) genes were inspected, and those showing a flat pattern of expression were removed. A substantial number of *Capsaspora* genes were retrieved if their expression pattern resembles those of other periodic genes previously identified. B: Ranking of JTK_CYCLE and RAIN (Hughes et al., 2010; Thaben and Westermark, 2014) scores for the new dataset. Scatter plot of ranks assigned by RAIN and JTK; each dot represents a *Capsaspora* transcript expressed throughout the cell cycle. Blue dots indicate transcripts retrieved in A. C: principal component and coordinate analysis reassuring that variance across samples is driven by periodic gene expression. D: heatmap showing gene expression across time for six largest clusters retrieved by hierarchical clustering. E: Barplots of Gene Ontology terms significantly enriched in for each cluster depicted in D. F: t-SNE plot of the periodic genes defined in A. G: pie chart showing the distribution of genes by gene age in the whole transcriptome of *Capsaspora* and in the dataset from A (up); and gene age enrichment for each of the clusters in D relative to the rest of *Capsaspora* genes (down).

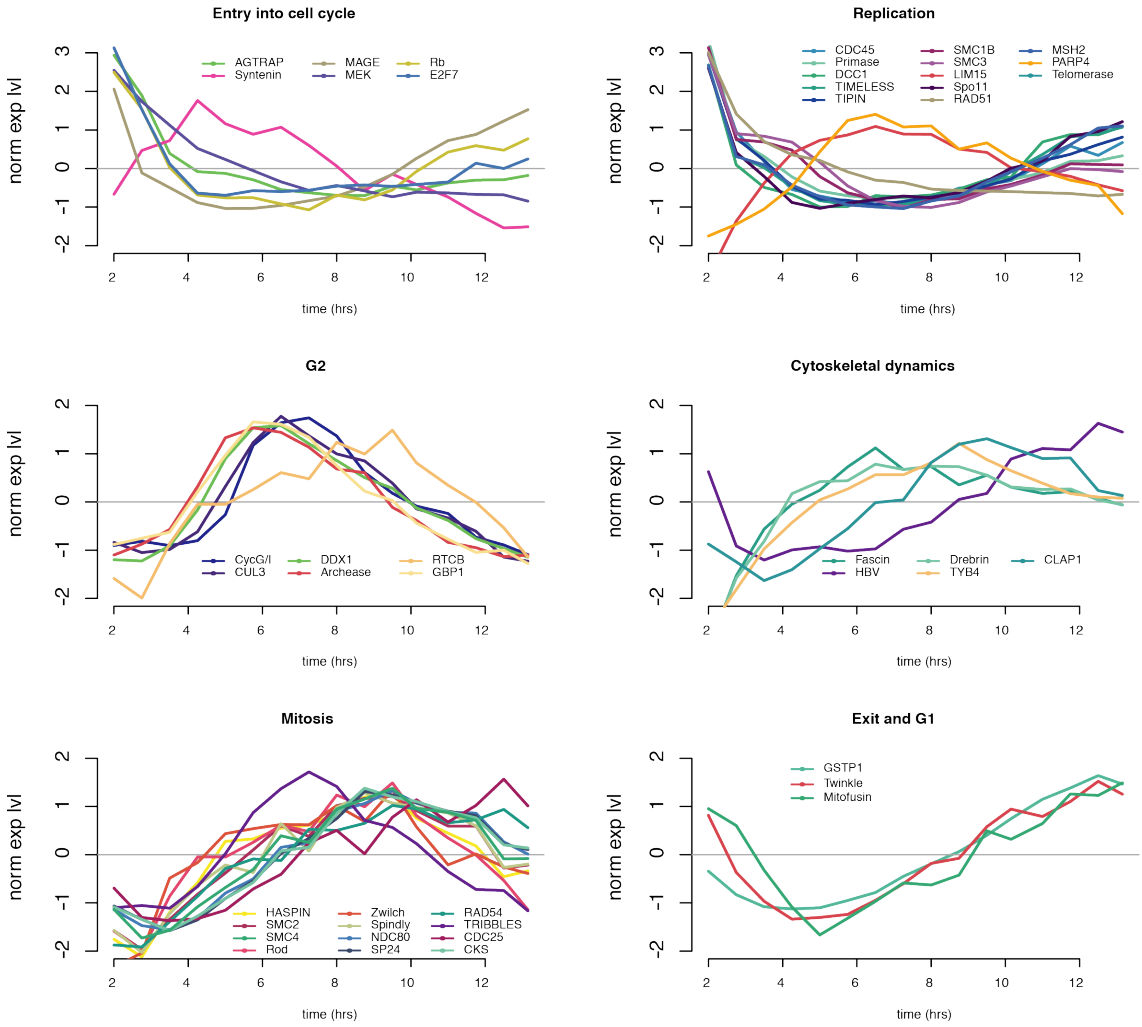


Figure A1.3: Expression of cell cycle regulators in *Capsaspora*. Expression profiles of different *Capsaspora* genes with known roles during the cell cycle in other species, grouped by their timing of expression and involvement in different processes throughout the cell cycle. Legends refer to the gene name assigned to their human orthologs.

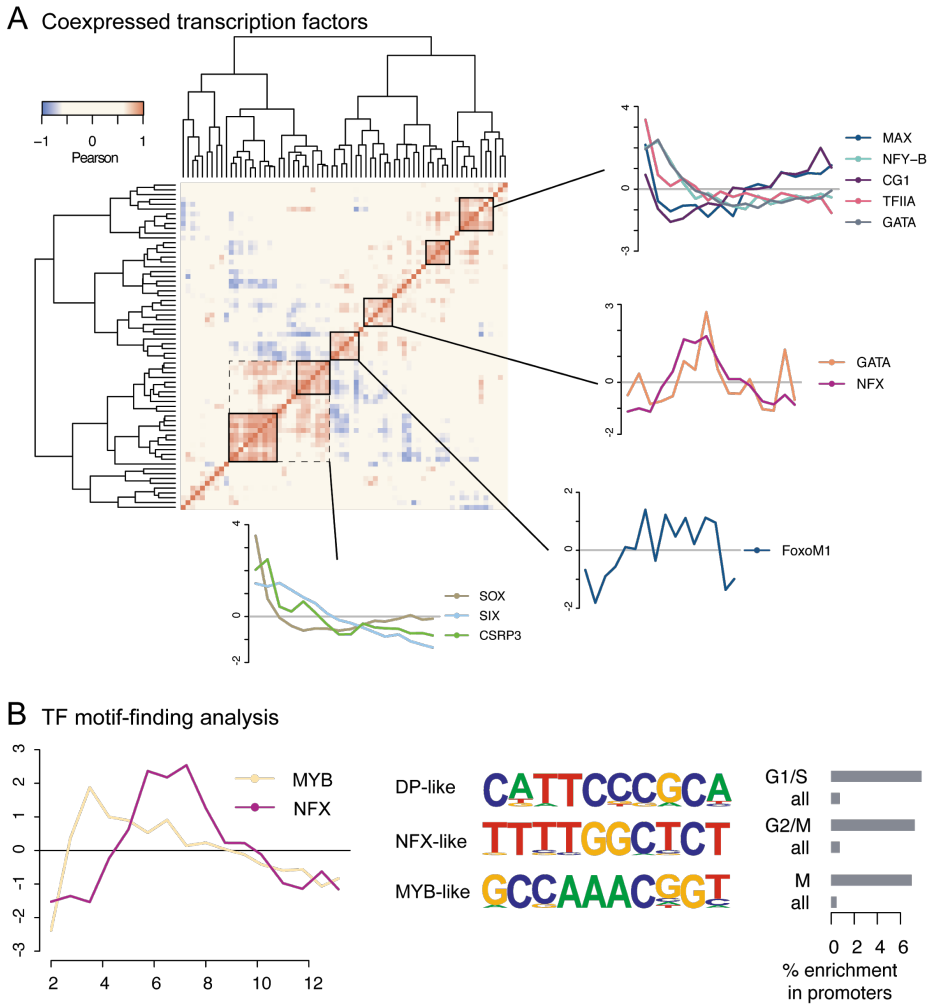


Figure A1.4: Exploring transcription factor dynamics during the cell cycle of *Capsaspora*. A: Pearson correlation between gene expression of transcription factors during the cell cycle of *Capsaspora*. Highlighted clusters contain genes with a similar expression pattern as those shown in the plots. B: Transcription factor motif analysis in the promoters of different cell cycle gene clusters in Capsaspora. Left: Expression pattern of *Capsaspora* transcription factors MYB and NFX; Right: significant motifs matching known cell cycle TFs or TFs periodically expressed in *Capsaspora*. Barplot shows percentage of enrichment in promoters of genes of a particular *Capsaspora* cell cycle cluster, relative to the whole number of promoters.

Table 1: all significantly enriched DNA motifs founds in the promoters of the different gene clusters in the periodic transcriptional program of *Capsaspora* (pval<10^-6).

G1/S

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details
1 *		1e-10	-2.405e+01	10.84%	1.31%	68.7bp (81.4bp)	skn-1/MA0547.1/Jaspar(0.720) More Information | Similar Motifs Found
2 *		1e-9	-2.083e+01	21.08%	6.47%	83.8bp (88.2bp)	dl/dmmpmm(Pollard)/fly(0.785) More Information | Similar Motifs Found
3 *		1e-8	-2.024e+01	7.23%	0.59%	73.1bp (96.5bp)	Dof2/MA0020.1/Jaspar(0.722) More Information | Similar Motifs Found
4 *		1e-8	-2.023e+01	7.83%	0.75%	85.2bp (88.4bp)	TFDP1/MA1122.1/Jaspar(0.797) More Information | Similar Motifs Found
5 *		1e-8	-1.972e+01	6.02%	0.35%	87.4bp (93.2bp)	Ct/dmmpmm(Noyes_hd)/fly(0.736) More Information | Similar Motifs Found
6 *		1e-8	-1.925e+01	11.45%	2.06%	75.0bp (87.4bp)	RIM101/MA0368.1/Jaspar(0.593) More Information | Similar Motifs Found
7 *		1e-7	-1.832e+01	8.43%	1.08%	89.2bp (77.3bp)	RUNX1(Runt)/Jurkat-RUNX1-ChIP-Seq(GSE29180)/Homer(0.725) More Information | Similar Motifs Found
8 *		1e-7	-1.819e+01	9.64%	1.51%	84.0bp (86.3bp)	RME1(MacIsaac)/Yeast(0.816) More Information | Similar Motifs Found
9 *		1e-7	-1.760e+01	6.02%	0.47%	85.2bp (92.1bp)	SAMD4A(SAM)/Homo_sapiens-RNCMPT00063-PBM/HughesRNA(0.724) More Information | Similar Motifs Found
10 *		1e-7	-1.674e+01	4.22%	0.15%	35.0bp (48.4bp)	AT5G02460(C2C2dof)/col-AT5G02460-DAP-Seq(GSE60143)/Homer(0.720) More Information | Similar Motifs Found
11 *		1e-7	-1.646e+01	3.61%	0.09%	77.8bp (75.2bp)	YY1(Zf)/Promoter/Homer(0.739) More Information | Similar Motifs Found

S

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details
1 *		1e-11	-2.614e+01	8.97%	1.07%	96.6bp (89.7bp)	DOF2(C2C2(Zn) Dof)/Zea mays/AthaMap(0.740) More Information | Similar Motifs Found
2 *		1e-11	-2.551e+01	3.59%	0.03%	74.5bp (81.9bp)	HNF4G/MA0484.1/Jaspar(0.632) More Information | Similar Motifs Found
3 *		1e-10	-2.423e+01	3.59%	0.05%	48.0bp (111.9bp)	Tb_0230(RRM)/Trypanosoma_brucei-RNCMPT00230-PBM/HughesRNA(0.674) More Information | Similar Motifs Found
4 *		1e-10	-2.373e+01	9.87%	1.57%	78.0bp (85.6bp)	Kr/dmmpmm(Papatsenko)/fly(0.681) More Information | Similar Motifs Found
5 *		1e-10	-2.353e+01	4.48%	0.15%	53.6bp (88.5bp)	Tv_0226(RRM)/Trichomonas_vaginalis-RNCMPT00226-PBM/HughesRNA(0.751) More Information | Similar Motifs Found
6 *		1e-10	-2.352e+01	19.28%	6.15%	76.0bp (86.2bp)	MYB88(MYB)/col-MYB88-DAP-Seq(GSE60143)/Homer(0.645) More Information | Similar Motifs Found
7 *		1e-10	-2.316e+01	3.59%	0.06%	76.6bp (62.1bp)	vnd/dmmpmm(Noyes_hd)/fly(0.762) More Information | Similar Motifs Found
8 *		1e-9	-2.249e+01	6.28%	0.53%	78.5bp (63.3bp)	OPI1/MA0349.1/Jaspar(0.680) More Information | Similar Motifs Found
9 *		1e-9	-2.247e+01	4.48%	0.17%	83.5bp (82.9bp)	ZAP1/MA0589.1/Jaspar(0.690) More Information | Similar Motifs Found
10 *		1e-9	-2.206e+01	3.14%	0.03%	80.7bp (36.0bp)	SF1(NR)/H295R-Nr5a1-ChIP-Seq(GSE44220)/Homer(0.647) More Information | Similar Motifs Found
11 *		1e-9	-2.176e+01	6.28%	0.57%	75.6bp (88.8bp)	HAP2/MA0313.1/Jaspar(0.708) More Information | Similar Motifs Found
12 *		1e-9	-2.140e+01	3.59%	0.08%	70.0bp (84.7bp)	MET28/MA0332.1/Jaspar(0.740) More Information | Similar Motifs Found
13 *		1e-9	-2.119e+01	23.32%	9.18%	89.6bp (87.6bp)	HNRNPA1L2(RRM)/Homo_sapiens-RNCMPT00023-PBM/HughesRNA(0.657) More Information | Similar Motifs Found
14 *		1e-8	-1.999e+01	3.59%	0.10%	79.0bp (93.5bp)	ABI4(1)(AP2/EREBP)/Zea mays/AthaMap(0.688) More Information | Similar Motifs Found
15 *		1e-8	-1.973e+01	11.21%	2.58%	67.6bp (84.3bp)	AT1G76870/MA1367.1/Jaspar(0.770) More Information | Similar Motifs Found
16 *		1e-7	-1.828e+01	3.14%	0.08%	76.7bp (91.4bp)	YPR022C/MA0436.1/Jaspar(0.722) More Information | Similar Motifs Found
17 *		1e-7	-1.769e+01	11.21%	2.89%	86.1bp (90.7bp)	MAC1(MacIsaac)/Yeast(0.750) More Information | Similar Motifs Found
18 *		1e-7	-1.649e+01	2.24%	0.02%	72.0bp (25.7bp)	At2g41835(C2H2)/col-At2g41835-DAP-Seq(GSE60143)/Homer(0.816) More Information | Similar Motifs Found

Table 1 (cont.):

G2/M

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details
1 *		1e-11	-2.625e+01	7.69%	0.70%	65.1bp (81.5bp)	elt-3/MA0542.1/Jaspar(0.734) More Information | Similar Motifs Found
2 *		1e-10	-2.357e+01	7.24%	0.73%	70.4bp (92.3bp)	NFIX/MA0671.1/Jaspar(0.675) More Information | Similar Motifs Found
3 *		1e-9	-2.272e+01	4.98%	0.24%	59.7bp (77.1bp)	OBP4(C2C2dof)/col-OBP4-DAP-Seq(GSE60143)/Homer(0.754) More Information | Similar Motifs Found
4 *		1e-9	-2.252e+01	4.07%	0.11%	60.1bp (79.5bp)	TDA9/MA0431.1/Jaspar(0.727) More Information | Similar Motifs Found
5 *		1e-9	-2.219e+01	5.88%	0.45%	77.3bp (83.1bp)	NFATC2/MA0152.1/Jaspar(0.730) More Information | Similar Motifs Found
6 *		1e-9	-2.170e+01	17.65%	5.56%	83.0bp (84.2bp)	SNRPA(RRM)/Homo_sapiens-RNCMPT00071-PBM/HughesRNA(0.692) More Information | Similar Motifs Found
7 *		1e-9	-2.154e+01	3.62%	0.08%	74.9bp (70.6bp)	ASD-1(RRM)/Caenorhabditis_elegans-RNCMPT00180-PBM/HughesRNA(0.695) More Information | Similar Motifs Found
8 *		1e-8	-2.054e+01	14.03%	3.81%	89.6bp (86.5bp)	ARNT::HIF1A/MA0259.1/Jaspar(0.691) More Information | Similar Motifs Found
9 *		1e-8	-1.957e+01	5.43%	0.46%	89.9bp (94.4bp)	MSN2/MSN2_H2O2Hi/1-MSN2(Harbison)/Yeast(0.862) More Information | Similar Motifs Found
10 *		1e-8	-1.862e+01	5.43%	0.51%	73.0bp (85.1bp)	KLF9/MA1107.1/Jaspar(0.728) More Information | Similar Motifs Found
11 *		1e-7	-1.838e+01	7.69%	1.25%	71.6bp (87.8bp)	CG11617/dmmpmm(Noyes_hd)/fly(0.751) More Information | Similar Motifs Found
12 *		1e-7	-1.835e+01	2.26%	0.01%	48.5bp (0.0bp)	MYB88(MYB)/col-MYB88-DAP-Seq(GSE60143)/Homer(0.676) More Information | Similar Motifs Found
13 *		1e-7	-1.713e+01	3.17%	0.11%	85.7bp (104.9bp)	PUCHI(AP2EREBP)/colamp-PUCHI-DAP-Seq(GSE60143)/Homer(0.729) More Information | Similar Motifs Found
14 *		1e-7	-1.677e+01	2.71%	0.05%	45.3bp (71.7bp)	MGP(C2H2)/colamp-MGP-DAP-Seq(GSE60143)/Homer(0.754) More Information | Similar Motifs Found
15 *		1e-7	-1.661e+01	7.24%	1.24%	85.1bp (82.5bp)	At3g60580(C2H2)/col-At3g60580-DAP-Seq(GSE60143)/Homer(0.715) More Information | Similar Motifs Found
16 *		1e-7	-1.618e+01	4.07%	0.29%	74.3bp (87.3bp)	RBP1(RRM)/Drosophila_melanogaster-RNCMPT00058-PBM/HughesRNA(0.716) More Information | Similar Motifs Found

M

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details
1 *		1e-11	-2.644e+01	9.45%	1.05%	72.4bp (93.1bp)	RFX1/MA0365.1/Jaspar(0.799) More Information | Similar Motifs Found
2 *		1e-11	-2.561e+01	6.97%	0.46%	80.9bp (73.3bp)	BMYB(HTH)/Hela-BMYB-ChIP-Seq(GSE27030)/Homer(0.705) More Information | Similar Motifs Found
3 *		1e-10	-2.308e+01	4.98%	0.19%	65.9bp (66.7bp)	KAN1/MA1027.1/Jaspar(0.784) More Information | Similar Motifs Found
4 *		1e-9	-2.190e+01	11.44%	2.16%	82.1bp (82.7bp)	HNRNPR(RRM)/Gallus_gallus-RNCMPT00288-PBM/HughesRNA(0.736) More Information | Similar Motifs Found
5 *		1e-9	-2.166e+01	3.48%	0.05%	63.8bp (101.0bp)	ZC3H10(Znf)/Homo_sapiens-RNCMPT00085-PBM/HughesRNA(0.647) More Information | Similar Motifs Found
6 *		1e-8	-2.068e+01	2.99%	0.02%	72.2bp (20.9bp)	RAR:RXR(NR),DR5/ES-RAR-ChIP-Seq(GSE56893)/Homer(0.660) More Information | Similar Motifs Found
7 *		1e-8	-1.941e+01	7.46%	0.95%	68.7bp (73.4bp)	PB0013.1_Eomes_1/Jaspar(0.767) More Information | Similar Motifs Found
8 *		1e-8	-1.941e+01	7.46%	0.95%	64.3bp (81.1bp)	SUP-12(RRM)/Caenorhabditis_elegans-RNCMPT00179-PBM/HughesRNA(0.795) More Information | Similar Motifs Found
9 *		1e-8	-1.880e+01	11.94%	2.81%	76.5bp (86.9bp)	IRF4(IRF)/GM12878-IRF4-ChIP-Seq(GSE32465)/Homer(0.705) More Information | Similar Motifs Found
10 *		1e-8	-1.862e+01	6.97%	0.85%	76.2bp (90.6bp)	At2g41835(C2H2)/col-At2g41835-DAP-Seq(GSE60143)/Homer(0.653) More Information | Similar Motifs Found
11 *		1e-7	-1.736e+01	22.39%	9.19%	82.9bp (80.1bp)	RBM28(RRM)/Homo_sapiens-RNCMPT00049-PBM/HughesRNA(0.693) More Information | Similar Motifs Found
12 *		1e-7	-1.624e+01	3.48%	0.14%	72.0bp (84.6bp)	XBP1/MA0414.1/Jaspar(0.696) More Information | Similar Motifs Found
13 *		1e-7	-1.621e+01	5.47%	0.58%	71.2bp (74.7bp)	STE12(MacIsaac)/Yeast(0.751) More Information | Similar Motifs Found

Table 1 (cont.):

	Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details
G1	1 *		1e-10	-2.400e+01	8.33%	0.90%	80.2bp (97.0bp)	HNRNPL(RRM)/Homo_sapiens-RNCMPT00091-PBM/HughesRNA(0.729) More Information | Similar Motifs Found
	2 *		1e-10	-2.366e+01	9.31%	1.23%	92.1bp (84.9bp)	ARF8/MA0944.1/Jaspar(0.721) More Information | Similar Motifs Found
	3 *		1e-9	-2.293e+01	9.80%	1.46%	89.0bp (87.0bp)	Eip74EF/dmmpmm(Bergman)/fly(0.710) More Information | Similar Motifs Found
	4 *		1e-9	-2.243e+01	2.94%	0.01%	111.9bp (0.0bp)	RAV1(1)(AP2/EREBP)/Arabidopsis thaliana/AthaMap(0.740) More Information | Similar Motifs Found
	5 *		1e-9	-2.207e+01	3.92%	0.07%	80.2bp (90.3bp)	Tv_0258(RRM)/Trichomonas_vaginalis-RNCMPT00258-PBM/HughesRNA(0.674) More Information | Similar Motifs Found
	6 *		1e-9	-2.133e+01	3.92%	0.09%	70.0bp (90.4bp)	Ot_0263(RRM)/Ostreococcus_tauri-RNCMPT00263-PBM/HughesRNA(0.678) More Information | Similar Motifs Found
	7 *		1e-8	-2.071e+01	6.86%	0.70%	79.5bp (75.3bp)	INO4/MA0322.1/Jaspar(0.790) More Information | Similar Motifs Found
	8 *		1e-8	-2.030e+01	7.35%	0.87%	101.3bp (96.6bp)	Tv_0258(RRM)/Trichomonas_vaginalis-RNCMPT00258-PBM/HughesRNA(0.700) More Information | Similar Motifs Found
	9 *		1e-8	-1.899e+01	4.90%	0.31%	84.6bp (78.3bp)	RDS2/MA0362.1/Jaspar(0.768) More Information | Similar Motifs Found
	10 *		1e-7	-1.832e+01	9.80%	1.95%	74.7bp (89.4bp)	AT5G25475(ABI3VP1)/col-AT5G25475-DAP-Seq(GSE60143)/Homer(0.667) More Information | Similar Motifs Found
	11 *		1e-7	-1.819e+01	3.43%	0.09%	83.3bp (67.8bp)	Sf3b4(RRM)/Danio_rerio-RNCMPT00224-PBM/HughesRNA(0.665) More Information | Similar Motifs Found
	12 *		1e-7	-1.780e+01	8.33%	1.42%	84.1bp (81.9bp)	Mecom/MA0029.1/Jaspar(0.770) More Information | Similar Motifs Found
	13 *		1e-7	-1.749e+01	9.80%	2.06%	100.1bp (85.7bp)	SD0002.1_at_AC_acceptor/Jaspar(0.675) More Information | Similar Motifs Found

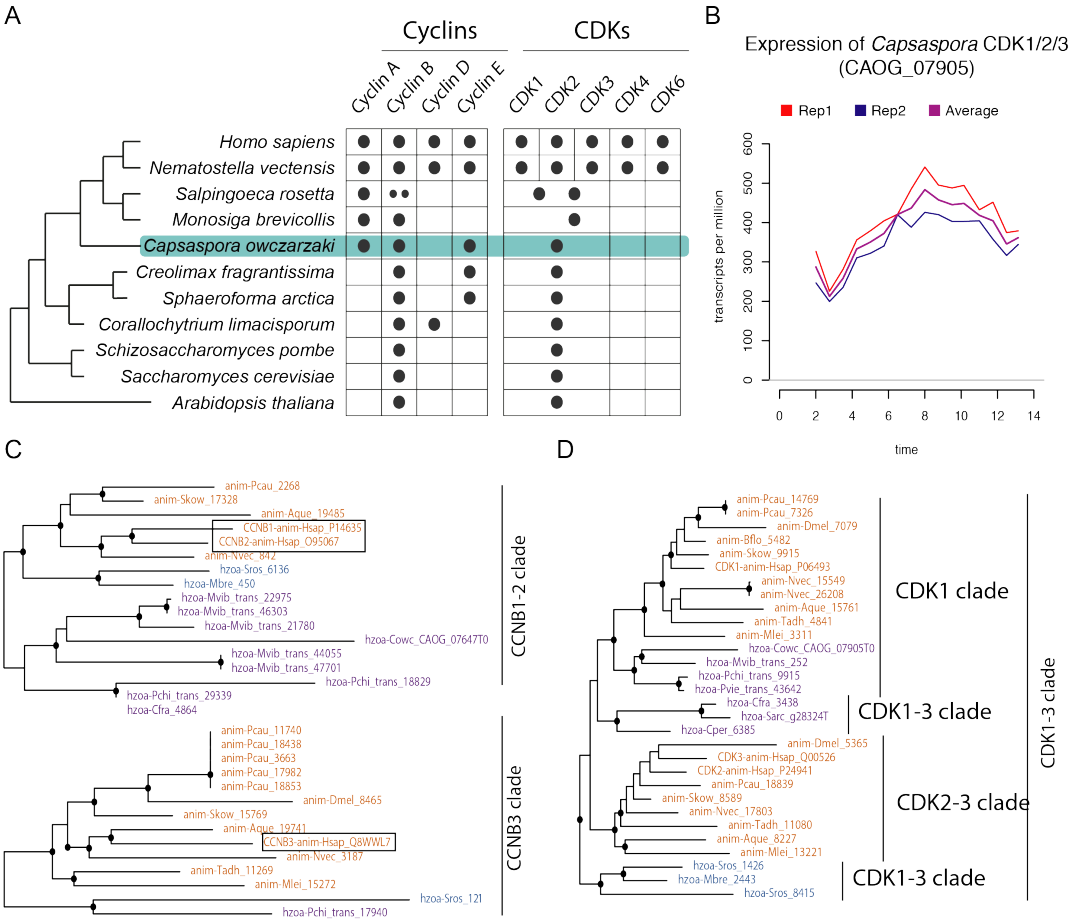


Figure A1.5: Phylogenetic analyses of the core cyclin-CDK system in opisthokonta, including unicellular relatives of animals. A: Presence and absence of the main animal drivers of the cell cycle. A simplified set of organisms is depicted (see Result 3.1 for more). B: Expression pattern of *Capsaspora* CDK1 during the cell cycle. C: zoom-in at the phylogenetic tree of cyclins found in Result 3.1. Orange represents animals, blue represents choanoflagellates, and purple the rest of unicellular relatives of animals used in the analysis. Species name follow the same four letter code as in Result 3.1. Dotted nodes indicate a Bootstrap higher than 90. D: Same as C for the phylogenetic tree of CDKs in animals and their unicellular relatives.

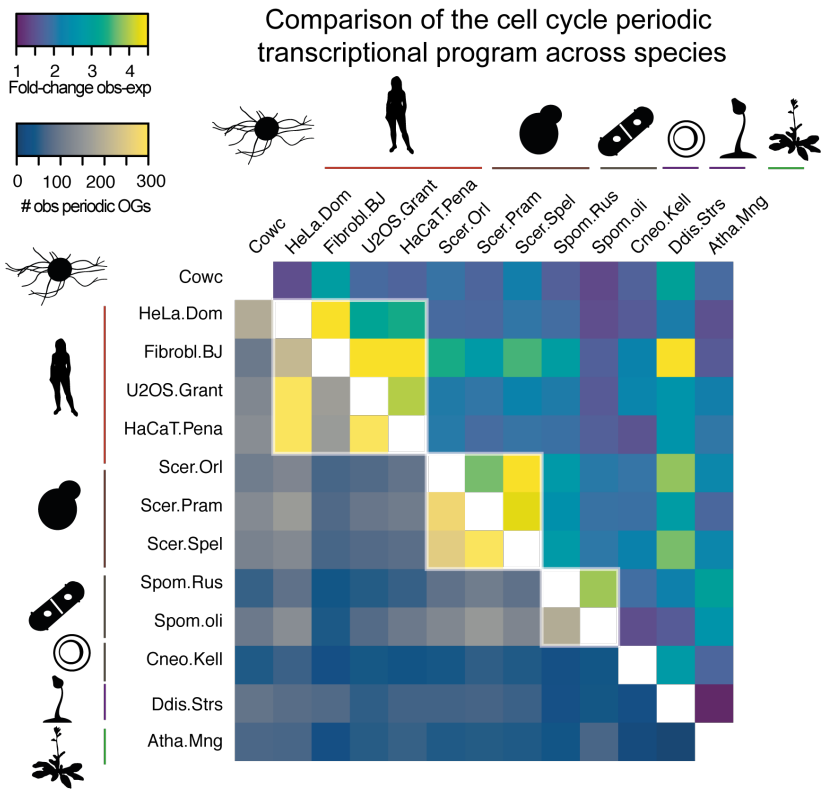


Figure A1.6: conservation of the periodic transcriptional program across species. Two overlapped heatmaps showing number of shared periodic orthogroups between pairs of species (lower left), and fold-change of expected vs observed numbers (upper right). Six additional datasets were added in this comparison: one additional human cell line, two budding yeast datasets, one fission yeast dataset, one dataset from *Cryptococcus neoformans* (Kelliher et al., 201_), and one dataset of E2F/Rb-regulated genes in *D. discoideum*.

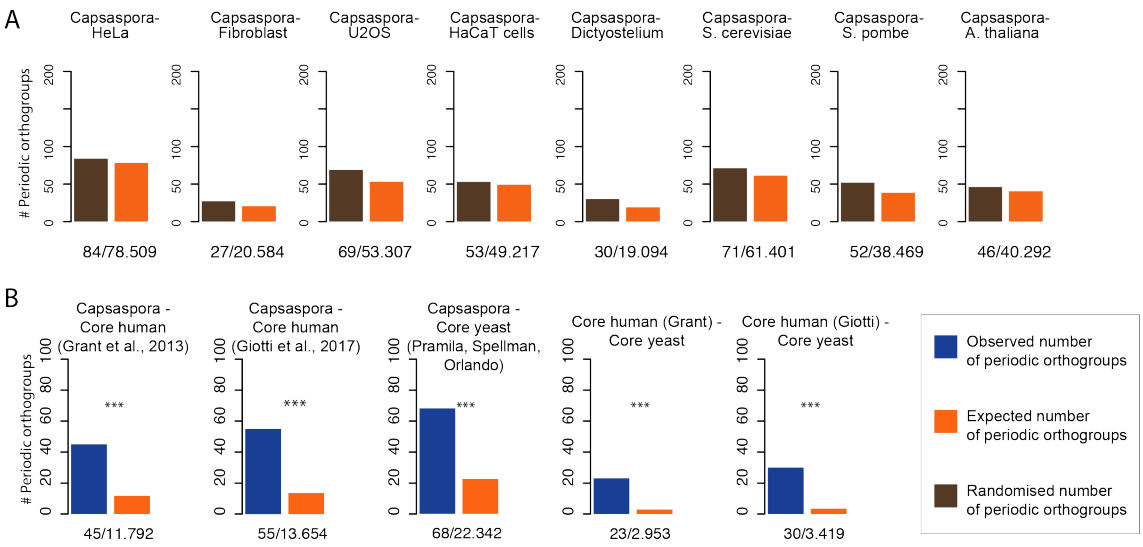


Figure A1.7: further investigation to comprehend similarities between periodic transcriptional programs. A: evaluation of phylogenetic closeness by random selection of orthogroups as gimmick-periodic. Fold-change (and hence statistical significance) drops compared to actual comparisons. B: evaluation of the cell cycle transcriptional program of *Capsaspora, Homo sapiens*, and *S. cerevisiae*. Three asterisks indicate p-value smaller than 10e-6.

Appendix 2: assessing cell viability upon DNA transfection

In the present thesis, I performed initial tests of a joined protocol of DNA transfection and culture synchronization[1]. Briefly, Capsaspora cultures were transfected as described in methods from Result 3.2 but, instead of left in fresh medium for eighteen hours, they were incubated in 10mM HU medium immediately after the glycerol shock, following the protocol of arrest described in Result 3.1. Preliminary results (data not shown) evidenced that, although yielding positive cells, our current protocol of DNA transfection may not be efficient enough to allow automated monitoring or quick screening of transfected cell throughout the cell cycle, unless coupled to more costly procedures such as FACS-assisted pooling of transfected cells. In order to generate homogeneous populations of cells with the same perturbation, stable transfection and genome-editing tools for Capsaspora have been in parallel development to the elaboration of a transfection protocol.

To develop genome-editing tools in Capsaspora, an initial protocol for genotyping was designed combining DNA transfection of fluorescence markers and genome editing constructs, cell sorting, PCR and Sanger sequencing, in order to obtain sequence information from populations of cells coming from an unique event of DNA transfection[2] (Fig. A2.1B). Transfected cells were collected in 96-well plates containing fresh medium; as a negative control, we also took cells that went under the transfection protocol but did not show fluorescence (putative negative cells), and cells that were grown under the same conditions but did not go through the transfection protocol (Fig. A2.1C).

We found that positive cells showed a much lower survival rate (7-13%) than the controls (~90%), measured as the fraction of thriving colonies observed through microscopy seven days after sorting per initial number of cells sorted (Fig. A2.1D,E). The low viability could not be attributed either to lethal mutations, as cells transfected only with the fluorescent marker also died, or to the simple effect of passing through the sorter

[1] People involved: Alberto Perez-Posada
[2] People involved: Alberto Perez-Posada, Andrej Ondracka, Helena Parra, Iñaki Ruiz-Trillo

or even a synergic effect of transfection and sorting, as negative cells survived at a much higher rate regardless of having passed through the transfection treatment. Upon further testing with different amounts of DNA, we observed that fluorescence intensity of positive cells was lower than that of cells transfected with two-to-ten times less amount of foreign DNA (Fig. A2.1F). This led us to hypothesize that high DNA-transfected cells suffered from a cytotoxic effect leading to a decay of cells expressing larger quantities of the protein (putatively by having uptaken larger amount of foreign DNA). This also matched parallel observations of Capsaspora cells transfected with the membrane marker having heavily branched filopodia, a phenotype not observed under normal culture conditions. DNA amount would then be crucial for the viability of cells, matching previous observation of DNA cytotoxicity in the literature affecting viability and transfection efficiency (Feng et al., 2007; Lesueur et al., 2016).

We hypothesize that part of the cytotoxic effect could be explained by the use of a promoter from a constitutively, very highly expressed gene, the Capsaspora elongation factor I (see Result 3.2). Further assays of viability confirmed that cells transfected with less amount of DNA showed a higher fluorescence intensity and had a survival rate of ~80%, comparable to that of negative cells. We thus discovered a cytotoxic effect on the viability of cells, and proceeded to adjust our protocol accordingly. As the duration of transient transfection was measured on cells with the initial amount of DNA, we speculate that the higher survival rate of low DNA-transfected cells would have an impact in the duration, possibly extending it. Future transfection assays in *Capsaspora* should also consider the use of low-expression cassettes based on our current transcriptomic knowledge of *Capsaspora* gene expression.

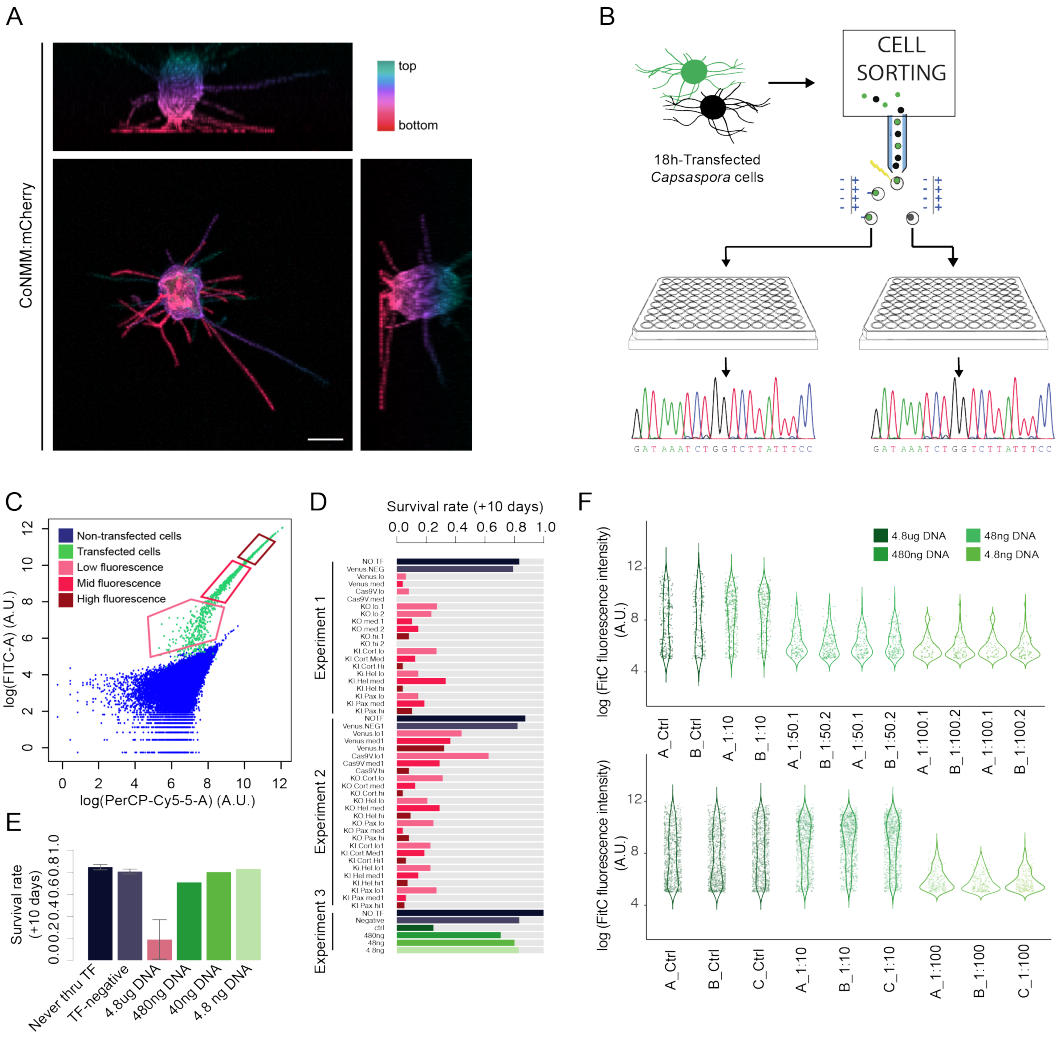


Figure A2.1: Unraveling cellular behavior and viability using transfection in Capsaspora. A: Our results revealed that Capsaspora cells use their filopodia to hold on to the substrate, illustrating the use of transfection as a tool to track cellular structures. B: An experimental setup designed to genotype Capsaspora cells upon transfection with genome-editing tools. Capsaspora cultures undergo transfection with a fluorescent marker and genome-editing constructs, and are passed through a cell sorting device to isolate individual transfected cells in 96-well plates. After a prudential time, say ten days, individual cultures are sampled for PCR and sequencing. C: Flow cytometry diagram showing a culture of transfected Capsaspora cells. Transfected cells show a wide range of fluorescence intensity, which can correlate with the amount of protein and with the amount of vector uptaken. Cells from each of the three gates were sampled to assess viability. D: Measurement of cell viability in all the different experiments. Survival rate is measured as the ratio of thriving colonies coming from an experiment as B, divided by the total number of cells isolated for each condition. Red colors indicate the same fluorescence intensity as shown in C. E: summarized plot of D, relative to the amount of DNA used in the transfection protocols. F: violin-plots showing the distribution of transfected cells by fluorescence intensity in each condition, for

two independent experiments. Colors indicate the amount of DNA used to transfect the cultures. A, B and C indicate biological replicates from independent cultures transfected the same day. In the first experiment, 1 and 2 indicate replicates coming from the same culture.

References:

Feng L, Guo M, Zhang Shuxiang, Chu J, Zhuang Y, Zhang Siliang. 2007. Optimization of transfection mediated by calcium phosphate for plasmid rAAV-LacZ (recombinant adeno-associated virus–β-galactosidase reporter gene) production in suspension-cultured HEK-293 (human embryonic kidney 293) cells. *Biotechnol Appl Biochem* **46**:127. doi:10.1042/ba20060143

Lesueur LL, Mir LM, André FM. 2016. Overcoming the Specific Toxicity of Large Plasmids Electrotransfer in Primary Cells In Vitro. *Mol Ther - Nucleic Acids* **5**:e291. doi:10.1038/mtna.2016.4