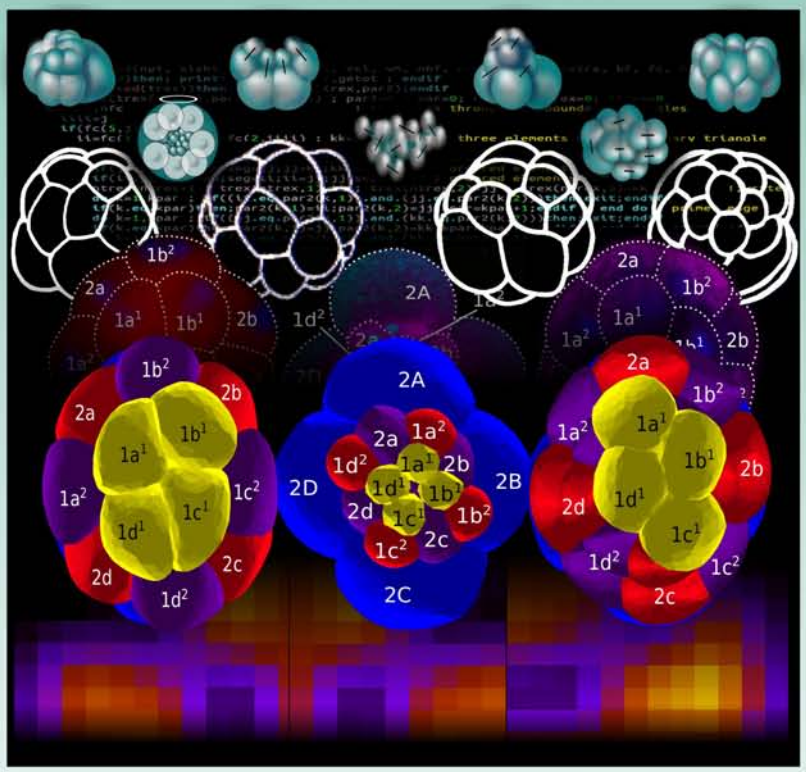
Universitat Autònoma de Barcelona

Computaional modeling of early cleavage in metazoans: the case of the spiral pattern

MIGUEL ANGEL BRUN USAN

Universitat Autònoma
de Barcelona
Programa de doctorado en Genética

Computational modeling of early cleavage in metazoans: the case of the spiral pattern



Doctoral thesis
(Tesis doctoral)

Miguel Angel Brun Usan
Director: Dr. Isaac Salazar Ciudad

2016

Universitat Autònoma de Barcelona

Computaional modeling of early cleavage in metazoans: the case of the spiral pattern

(Modelado computacional del clivaje temprano en metazoos: el caso del patrón espiral)

– DOCTORAL THESIS --

Memoria para optar al grado de Doctor por la Universitat Autònoma de Barcelona. Programa de Doctorado en Genética (Departament de Genètica i microbiologia. Grup de Genòmica, Bioinformàtica i evolució)

Author: MIGUEL ANGEL BRUN USAN **Director: Dr. Isaac Salazar Ciudad**

Bellaterra, de de 2016.

Facultad de Biociències

Departament de Genètica i de Microbiologia
(Grup de Genòmica, Bioinformàtica i Evolució)
Facultat de Biociències. Edifici C
08193 Bellaterra (Cerdanyola del Vallès)
TEL +34 93 581 27 24
FAX +34 93 581 23 87

Dr. ISAAC SALAZAR CIUDAD, Profesor Contratado Doctor y director de la presente Tesis doctoral,

HACE CONSTAR QUE:

La Tesis escrita por Miguel Angel Brun Usan es apta para ser leída y defendida en público, puesto que reúne todos los requisitos propios de este tipo de trabajo: rigor científico, resultados originales y aplicación de una metodología adecuada. No menos importante, la linea de trabajo desarrollada en esta Tesis es fácilmente generalizable a otros grupos animales, abriendo nuevas vías de investigación futura. Además, el doctorando ha demostrado capacidad y autonomía a la hora de implementar las diferentes lineas de investigación.

La presente Tesis cumple con creces los objetivos planteados en el proyecto inicial. Desde el principio, se propuso crear un modelo computacional capaz de reproducir los diferentes patrones de clivaje observados durante el desarrollo temprano de diferentes animales. Dicho modelo permitiría arrojar luz sobre los diferentes procesos celulares que dan lugar a los diferentes patrones de clivaje, así como sobre los factores que median en las transiciones evolutivas entre esos patrones. Además, contando con un modelo suficientemente realista, los resultados obtenidos podrían compararse cuantitativamente con embriones reales.

La complejidad técnica del proyecto exigió un planteamiento secuencial en la ejecución del mismo: 1) Con el objetivo de introducir al doctorando en la modelizacion bioinformática, se elaboró un modelo

computacional (*toy-model*) sobre la evolución adaptativa de fenotipos complejos.

2) Se desarrolló un modelo matemático general de desarrollo embrionario capaz de simular el desarrollo de cualquier tipo de órgano o tejido. El doctorando, utilizando las habilidades informáticas adquiridas en el punto (1), participó desarrollando partes relevantes que permitirían mas tarde el modelado de los patrones de clivaje.

3) Se llevó a cabo un trabajo de revisión bibliografica que resume el conocimiento actual en el desarrollo temprano de todos los grupos animales, haciendo hicapié en los mecanismos celulares implicados en los diferentes patrones de clivaje.

4) Con el modelo desarrollado en el punto (2) y el conocimiento adquirido en el punto (3), se elaboró un modelo de clivaje espiral. El patrón de clivaje espiral es un tipo de clivaje representativo y extendido entre los metazoos y del que se dispone de abundante bibliografía al respecto. Ademas, la elección del patrón espiral permitió contar con embriones reales que comparar con los resultados.

Cabe destacar que cada fase del proyecto ha dado lugar a un artículo científico en publicaciones de prestigio internacional y considerable impacto (*Journal of Evolutionary Biology (1), Bioinformatics (2), Development (4)*) (A fecha de depósito de Tesis, 27 de Septiembre de 2016, la última se halla aceptada pero pendiente de publicación)

Los resultados presentados en esta tesis contribuyen de manera significativa al avance del campo de la biología evolutiva y de la biología del desarrollo. Por un lado, los experimentos *in silico* realizados con el modelo de desarrollo temprano han arrojado luz sobre el largamente debatido origen del patron de clivaje espiral, y han permitido rechazar algunas teorías preexistentes al respecto. Por otro lado, las nuevas técnicas informáticas y de morfología cuantitativa desarrolladas en la presente tesis son de amplia aplicabilidad dentro del campo del desarrollo temprano, por lo que otros grupos de investigación podrían beneficiarse de las mismas.

En general, la calidad de esta tesis es alta, tanto en lo referente al número de publicaciones como a la calidad de éstas. Durante la elaboración de esta tesis, el doctorando ha demostrado una comprensión profunda en el campo de la biología evolutiva y de la biología del desarrollo, así como una capacidad de sintetizar dicha comprensión en modelos matemáticos que a buen seguro tendrán un impacto duradero en el campo.

Dr. Isaac Salazar Ciudad. En Barcelona, a de de 2016

FUNDING:

During most of time (year 2011 to 2015), this Thesis was funded by the Spanish Ministry of Science and Innovation by means of the Grant (BES2011-046641), framed within the project (BFU2010-17044: “Modelos de redes genéticas y celulares en el desarrollo embrionario animal y la evolución”). In 2016, it was partially funded by the the Finnish Academy (WBS 1250271) and partially by the Spanish INEM.

To all people who have taught me new things

(A todos los que me han enseñado cosas nuevas)

ACKNOWLEDGEMENTS:

I thank my supervisor Isaac Salazar-Ciudad for his advice and training, and for allowing me to enter in the world of science. Other scientists have been fundamental during the development of this Thesis: Dr. Jukka Jernvall, from Helsinki University in Finland, Dr. Stuart A. Newman, from the New-York Medical College in USA and Dra. Cristina Grande, from the CBMSO, in Madrid. Thank you. I also thank my Lab colleagues and coworkers: Miquel Marin-Riera, Irepan Salvador-Martinez, Roland Zimm, Tommi Välikangas, Anton Chernenko, Fernando Urbano, Pascal Hagolani, Alexis Matamoro and Mona Christensen for their helpful comments, fruitful discussions, manuscript reviews and for creating a friendly and peaceful working environment.

I would like also thank the teachers and researchers of the Departament de Genetica i microbiologia in Barcelona, who share with me their pasion for evolutionary biology. Special mention to Dr. Antonio Barbadilla and Dr. Alfredo Ruiz for their wise advices. I also thank the members of the Thesis Monitoring Committee, Dra. Alba Hernandez, Dr. Ricard Marcos and Dr. Alfredo Ruiz for their advices and guidance. I am also very thankful to the Thesis Tribunal, and to all the presents for your assistance.

In addition, I would like also mention some people who, not strictly belonging to the academic area in which this Thesis has been developed, have been fundamental in order for it to be carried out. I would like to thank them. My parents, for the freedom of choice they offered me and their unconditional support. My brother, for becoming an example of personal overcoming. To my uncle Jose Luis Usan and my grandfather Jose Usan, since they let me know the pleasure of the knowledge of nature. To my grandmother, for her long-distance support. Also to my uncle Jaime Brun for his “inverse motivation” (he told me during my childhood I will never become a biologist). To Cris, who endures the science as my other love, and keeps me firmly attached to the solid ground of tangible things. Thanks.

During my learning stage of life (which lasts until now), many people have influenced me to be a biologist, and this I would like to thank them. They are the irreplaceable friends from my village (especially those belonging to the biological association SAB); my “grandfiend” Malocha, authodidact naturalist; and the adventurers Jesus and Lucía, who taught me that knowledge does not finish just in books. Some wonderful teachers I have had in the public Spanish education and academic system also deserve a place in this list: Ruben Peña, Jesus Molledo, Carmelo, Manuel Hernandez, Patricio Dominguez … and DonGuillermo (written that way), whose lessons were well

beyond the scope of primary school.

Many biologists and scientists have also contributed to sculpt my passion for biology and science in general, and thus have played a role in this thesis, as large-effect motivational agents: Mario, Isaac, Irene, Agus, Freiji, Ali, Pantxo, Yulan, Gabi, Gise, Luz, Carmen, Wessos, Paula, Joaquin, Ruben, Whyshu, Marco, the physicists Ruben and Daniela … (the list is open ended and always growing). Thank you all for your almost infinite talks about evolution, cosmology, nature, ethology, sociology, physic, maths …. and all those marvelous topics contained within this Universe.

Despite the fact I do not know them personally (some of them rest at just one node far away in the dense network of human and academic relationships), some people have influenced my scientific thought such a way I feel obligated to include them in this acknowledgment list: the more prominent are Isaac Asimov, Carl Sagan, Stephen Jay Gould, Clifford Pickover, Gerald Durrell, David Attenborough, Stuart Kauffman, Brian Goodwin and Lynn Margulis. Thank for your legacy.

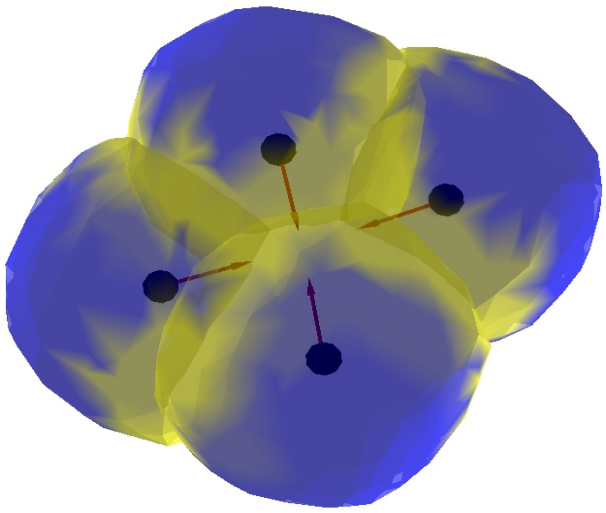


TABLE OF CONTENTS:

UNIVERSITAT AUTÒNOMA DE BARCELONA

SUMMARY

Facultat de Biociències

Departament de Genètica i Microbiologia

Doctor of Philosophy

Computational modeling of early cleavage in metazoans: the case of the spiral pattern

by Miguel Angel BRUN USAN

Cleavage is the earliest developmental stage. In it, the fertilized oocyte becomes partitioned into a set of smaller cells (blastomeres) with a particular spatial arrangement (a cleavage pattern). Different metazoan species have different cleavage patterns, but all of them fit into a small set of basic types according to some geometrical regularities of their blastomere arrangements.

Distantly related taxa can exhibit similar cleavage patterns, while species belonging to the same taxon can have different cleavage types, resulting in a weak relationship between cleavage patterns and phylogeny. These discrepancies between the cleavage patterns and the phylogenetic position of metazoan taxa are difficult to explain and seem very counterintuitive unless we gain more knowledge on how the spatio-temporal combination of different cell processes can generate the different cleavage patterns.

Taking this into account, the aim of this Thesis is twofold:

First is to gain an overview of these cell processes involved in metazoan cleavage, and to review how they are combined in the early development of the major metazoan groups. We did it by means of a thorough bibliographic research, which shows that most taxa use the same basic set of cell processes (such as directed cell division or cell adhesion) to build their cleavage patterns. Most of these processes are evolutionary old, and involve a series of epigenetic factors (such as physical interactions between neighboring cells, or between these cells and the outer envelope of the embryo), that are not encoded as such in genes.

The second aim of this Thesis is to explore how these basic cell processes can account for one of the most widespread cleavage pattern: the spiral pattern. Spiral cleavage is the most abundant cleavage type at the phylum level, but there is no consensus about which are the cellular processes responsible of the specific spatial arrangement of blastomeres in the spiral blastula. In order to do this, we use a new computational model of cell and tissue bio-mechanics to implement the different existing hypotheses about how the specific spatial arrangement of cells in spiral cleavage arises

during development. By means of this model, we found that cell polarization by an animal-vegetal gradient, Sachs' rule (a bias towards perpendicularity between consecutive cell divisions), cortical rotation and cell adhesion, when combined, reproduce the spiral cleavage while other combinations of processes can not. In addition, we reproduce the cell spatial arrangement of the blastulae of seven different species (four snails, two polychaetes and a nemertean), which suggests that the morphological variation observed in spiralian cleavage can be explained by quantitative variations in the relative strength of each of these few cell processes.

Concerning the distribution of the different species analyzed in the parameter space, we found that some spiralian cleavage patterns can be produced by many different combinations of parameter values (they are found in a large volume of the parameter space) while others require a much more restricted combination of parameter values. We also find that the spiral cleavage patterns of species that are phylogenetically closer are not necessarily closer to each other in the parameter space than to the patterns of other, more phylogenetically distant, species. This suggests that the relationship between the underlying developmental parameters (which are relatively easy to change by mutational changes) and the resulting cleavage patterns is not simple.

Finally, we discuss whether these findings about the spiral cleavage pattern, which are in agreement with our current understanding of general development dynamics, can be generalized to other non-spiralian systems.

UNIVERSITAT AUTÒNOMA DE BARCELONA

RESUMEN

Facultat de Biociències

Departament de Genètica i Microbiologia

Doctor

Modelado computacional del clivaje temprano en metazoos: el caso del patrón espiral

de Miguel Angel BRUN USAN

El clivaje es la primera etapa del desarrollo embrionario. En ella, el zigoto fecundado se divide en un conjunto de células menores (blastómeros), que se disponen en una organización espacial concreta (un patrón de clivaje). Diferentes metazoos poseen diferentes patrones de clivaje, pero todos ellos pueden ser clasificados en un pequeño conjunto de tipos básicos atendiendo a algunas regularidades geométricas en la disposición de sus blastómeros.

Taxones con un parentesco evolutivo lejano pueden presentar patrones de clivaje similares, mientras que especies pertenecientes a un mismo taxón pueden tener patrones de clivaje diferentes, lo cual resulta en una correlación escasa entre los patrones de clivaje y la filogenia. Estas discrepancias entre los patrones de clivaje y la posición filogenética de los taxones de metazoos es difícil de explicar, y resulta contraintuitiva a menos que se conozca cómo la combinación espaciotemporal de diferentes procesos celulares puede dar lugar a los diferentes patrones de clivaje.

Teniendo esto en cuenta, el objetivo de esta Tesis es doble:

El primer objetivo es aportar una visión global de los principales procesos celulares implicados en el clivaje de metazoos, y revisar cómo éstos se combinan durante el desarrollo temprano de los principales grupos de metazoos. Esta parte se llevó a cabo mediante una revisión biblográfica detallada, la cual muestra que la mayoría de los taxones utilizan el mismo conjunto de procesos celulares básicos (como la división celular dirigida, o la adhesión celular) para generar sus respectivos patrones de clivaje. La mayoría de estos procesos son evolutivamente antiguos, e involucran una serie de factores epigenéticos (como las interacioens físicas entre células adyacentes, o entre éstas y la cubierta externa del embrión), que no están codificados genéticamente como tales.

El segundo objetivo de ésta Tesis es explorar como éstos procesos celulares básicos estan implicados en uno de los patrones de clivaje mas extendidos: el patrón espiral. El patrón de clivaje

espiral es el tipo de clivaje mas abundante a nivel de phylum, pero no hay consenso sobre cuales son los procesos celulares responsables de la distribución espacial de los blastómeros típica de las blástulas espirales. Para llevar esto a cabo, hemos utilizado un nuevo modelo computacional de la biomecánica de células y tejidos, en el que hemos implementado las diferentes hipótesis existentes sobre la emergencia del clivaje espiral mediante diferentes procesos celulares. Mediante este modelo, encontramos que la combinación de una polarización celular por un gradiente animal-vegetal, la regla de Sachs (una tendencia a la perpendicularidad entre divisiones celulares sucesivas), rotación cortical y adhesión celular, reproducen el patrón espiral, mientras que otras combinaciones de procesos celulares no lo consiguen. Además, reproducimos la distribución espacial de blastómeros en las blástulas de siete especies diferentes (cuatro caracoles, un gusano poliqueto y un nemertino), lo que sugiere que la variación morfológica observada en el clivaje espiral puede explicarse por variaciones cuantitativas en la intensidad relativa de cada uno de los procesos celulares implicados.

En cuanto a la distribución de las diferentes especies analizadas en el espacio de parámetros, se encuentra que algunos patrones pueden producirse por medio de muchas combinaciones diferentes de parámetros (se encuentran en un gran volúmen del espacio de parámetros), mientras que otros requieren una combinacion de parámetros mucho mas restringida. También hallamos que los patrones espirales de especies que estan filogenéticamente próximas no estan necesariamente mas cerca en el espacio de parámetros de lo que lo están entre sí especies filogenéticamente distantes. Esto sugiere que la relación entre los parámetros de desarrollo subyacentes (que son relativamente fáciles de cambiar por medio de mutaciones genéticas) y los patrones de clivaje resultantes no es simple.

Por último, se discute si estos resultados relativos al patrón de clivaje espiral, que están en concordancia con nuestro conocimeinto actual de la dinámica general del desarrollo, pueden hacerse extensivos a otros tipos de clivaje no espiral.

INTRODUCTION:

The conceptual framework of this Thesis: Evo-devo. A brief historical background.

Before the dawn of evolutionary thought, the astonishing diversity of living forms was believed to be the manifestation of a divine plan (Paley, 1802). According to this pre-evolutionary way of thinking, each trait of an organism had been superbly designed by a creator to optimally perform a function (Gould and Lewontin, 1977), and thus the ultimate cause of each one of organs and morphologies found in nature was then reduced to be the divine intervention *ab initio*.

In the XIX century, the concept of natural selection developed by Darwin became the logical mechanism responsible for explaining the adequacy of each body part (organs, morphological structures) to the environment. According to Darwin, organisms change their phenotype (their set of traits) over generations by means of the iteration of a mechanism called natural selection. Basically, evolution by natural selection between two generation has three requirements (Darwin, 1859):

1. Variation in traits within a population (i.e., not all individuals are identical)
2. The inheritance of some of these trait variants (i.e., parents can “pass” the traits' variations to their offspring, and, thus, the offspring resemble their parents).
3. Ecological factors that lead to some heritable trait variants to be dis-proportionally transmitted to the next generation (i.e. Not all individuals have the same chance to survive (fitness), and the selected trait has to have a role on this probability).

Since its discovery, natural selection has been seen as the primary explanation for adaptive evolution. However, the “unity of type”, that refers to the fact that besides the adaptive part of the phenotype, phylogenetically related organisms (of the same “type”) share a common, not necessarily adaptive parts of the phenotype, was not easily explained in terms on natural selection alone. Thus, Darwin also recognized a second relevant evolutionary factor that would account for the unity of type. This second factor would be “common development” or embryonic homology (Darwin, 1959; Gilbert, 2003). That way, embryonic resemblances between two different animal groups would be a strong argument in favor of their evolutionary relatedness. This is, having a similar development would suggest that these two animal groups are related by common descent (Ospovat, 1981), what is especially useful in the cases where adaptive modifications make the comparison between adult's traits unfeasible (e.g. those derived from a parasitic lifestyle). Embryonic homology allowed early embryologist to classify correctly the barnacles within crustaceans, and ascidians within chordates (they were hitherto classified as mollusks), by establishing parallelisms between embryonic structures that might no longer be visible in the adult

organism (Darwin, 1859; Kowalewsky, 1866; Balfour, 1880; Bowler, 1996).

These achievements made embryonic development to be acknowledged as a relevant factor in evolution by many early evolutionary biologists (including Darwin himself). But these embryologists were mainly focused in using the embryonic homologies for establishing the phylogenetic relationship between animal groups (Muller, 1894; Hall 2000); or to disentangle the nature of the apparent parallelism between their embryonic stages and their evolutionary ancestors (Haeckel, 1866; Gould, 1977; Løvtrup, 1978). Only very few of the nineteenth-century evolutionary biologists were concerned with the potential role of the of the embryonic development in determining, along with natural selection, the direction of evolutionary change (Todes, 1989, Kropotkin, 2012).

By the end of the XIX century, the lack of new relevant empirical data and new conceptual frameworks (beyond mere speculation) caused embryology to become and independent field of evolutionary biology (Roux, 1894). According to Roux, embryology had to leave the seashore and forests and go to the laboratory, in order to gain a better understanding of the developmental dynamics *per se* (ibid). This new approach, characterized by a strong descriptive and empirical methodology, and by a marginal evolutionary perspective was known as *Entwicklungsmechanik* (developmental mechanics), and flourished in the late XIX century and in the early decades of the XX century (Gilbert and Raunio, 1997; Gilbert, 2003). In addition, the knowledge about the different developmental mechanisms was too immature as to understand how they may produce the different phenotypic variations, which in turn may determine the direction of evolutionary change.

Thus, when the modern evolutionary synthesis was built in the 1930s, the previously held notion that embryology may contribute to our understanding of the mechanism of evolution was left behind. For instance, in 1894, Bateson (one of the new synthesis' advocates) claimed that «the embryological method has failed» when it came to determining the mechanisms of evolution (Gilbert, 2003).

Instead, the new field of genetics (which emerged after the re-discovery of the Mendel laws by Correns and De-Vries in the year 1900) was to substitute embryology (and morphology) as the main source of evidence in evolutionary biology (Stomps, 1954). In fact, Morgan (1932) made the case for genetics being the sole scientifically valid approach to study of evolution (Gilbert, 2003). In the so called Evolutionary Synthesis (Morgan, 1932; Dobzhansky, 1937), the main driver of evolutionary change was natural selection, and individuals (including their phenotype) were merely reduced to genes (Weissman, 1875). Since these genes could change in whichever conceivable manner due to random genetic mutations, it followed that any transformation in the phenotype was possible (as long as this transformation is relatively small). This gene-centered view of evolution

was to continue as the mainstream evolutionary thought during most part of the XX century, and is exemplified in the much later works of R. Dawkins: “The fundamental units of natural selection, the basic things that survive or fail to survive, that form lineages of identical copies with occasional random mutations, are called replicators. DNA molecules are replicators”. Under this point of view, the organisms no longer existed as such, so that the morphology had only a secondary interest, as so had the embryonic development that created that morphology (Dawkins, 1976; Goodwin, 1996).

However, since the origin of the Evolutionary Synthesis, some authors warned about the idea of decomposing the whole organisms in a large set of independent traits. For instance, R. Fisher developed a conceptual model (known as Fisher's geometrical model) which shows that the more independent traits an organism has the less the likelihood of large-effect mutations to be adaptive (Fisher, 1930).

In the Annex I, we present a new computational model to explore how adaptation depends on the number of traits. Our model uses a Monte-Carlo experiments and explicit fitness functions rather than geometrical considerations (Brun-Usan et al., 2014), so that, unlike the Fisher's model, yields quantitative results. These results, that are in agreement with previous works (Martin and Leonormand, 2006; Lourenço et al, 2011), show that, irrespective the fitness function used, there are a rather low number of independent traits that can be effectively driven close to the optimum by natural selection (the optimum is an arbitrary fixed trait value, for example, a limb length of one centimeter). The percentage of traits that show adaptive change decreases as the mutation rate increases or the population size decreases, being the rest of the traits drifting away from the optimum towards non-adaptive values. This lack of optimality happens even in the best conceivable scenario for natural selection to act: independence between traits, nearly optimal initial conditions and realistic population sizes. Three new lines of inquiry not present in previous related works were introduced in this one: 1) an estimation of the maximum number of adaptive traits in a number of real model species (based on empirical data extracted from bibliography); 2) a mathematical-analytical derivation of the model and 3) an extension of the model in which the number of traits borne by the organisms are also subject to change and can be selected. This latter approach suggests that despite simple organisms start increasing adaptively the number of traits they bear, the selective advantage of having more traits disappears soon as the number of traits increase over evolutionary time. In addition, traits that appeared early in evolution tend to be closer to the optimum than the later ones.

Continuing with the historical background, in the 1940's, some researchers , such as Wright (1949) and Goldschmidt (1940), made attempts to integrate some developmental phenomena into

the Modern synthesis (also known as Neo-Darwinism). In the 1950's Waddington posed a more thorough critique to the Neo-Darwinism, based on three phenomena that it was unable to explain: 1) much variation appear to be non-genetic and regulated by the environment, not by the inherited genotype; 2) large groups of animals differ from each other in dramatic ways, not compatible with the piecemeal and continuous divergence in the traits from one to another (what Goldschmidt (1940) called «unbridgeable gaps») and 3) the palentological record shows different rates of evolution for different lineages, epochs and organs, with many morphologies which suddenly appear in the fossil record and then remain practically unchanged during vast periods of time (a phenomenon known today as punctuated equilibrium (Gould and Eldredge, 1977)). To solve these caveats, Waddington argued that a study of those processes that get the genotype from the phenotype was required: “Changes in genotypes only have ostensible effects in evolution if they bring with them alterations in the epigenetic processes by which phenotypes come into being; the kinds of change possible in the adult form of an animal are limited to the possible alterations in the epigenetic system by which it is produced» (Waddington, 1942). That is: natural selection is only concerned with phenotypes, not genotypes as such, and since phenotypes are built by development, one must understand the process of development if he or she is to understand evolution.

In addition, there were many biologists working in the the emerging field of the phenotypic plasticity: the dependence of the phenotypic outcome of development on the environment. Phenotypic plasticity also challenged the predominant gene-centric view because all the phenotypes that appeared under different environmental conditions (constituting a “reaction norm”) had the same genetic basis, so the relationship between the genotype and the phenotype is necessarily not simple (Waddington, 1942; Jablonka and Lamb, 2007).

The mainstream evolutionary biology continued, however, with a mostly reductionist gene-centric view. That view is was partially justified by the outstanding advances in the fields of genetics: the discovery of the DNA molecule as the carrier of genetic information (Avery et al., 1944), its stoichiometric dynamics (Chargaff, 1950) and molecular structure (Watson and Crick, 1953), the discovery of the genetic code (Niremberg et al., 1965) or the recombinant DNA techniques, just to cite a few. In addition, the robust mathematical treatment of the population genetics (Hardy, 1908; Lynch and Walsh, 1998) made possible to predict some results of the breeders' experiments, so that the fields of population genetics and quantitative genetics blossomed (Goodwin, 1996; Gilbert, 2003), and the emphasis was shifted from how genes act to build the phenotype to the rules of their segregation and transmission across generations and populations (Goodwin, 1996).

However, this long-lasting perception that evolution was mainly concerned with genes (with development playing no major role), relies on some assumptions that are rarely, or ever, stated explicitly. They are, however, logically required for natural selection to be the only most important force determining the direction of evolutionary change in the phenotype (Salazar-Ciudad, 2006). These assumptions are not always compatible with the current understanding of embryonic development. These assumptions can be summarized as follows: (1) there was additive genetic variation for most traits (Barton and Turelli, 1989); (2) the relationship between the phenotype and the genotype is simple and (3) morphological variation is gradual. Let us take a brief on their main oversights:

First, the argument that there is genetic variation for most traits is based on univariate traits (that is unidimensional traits such as body size or length) (Wijngaarden et al., 2002; Bronikowski et al., 2004). However, that a wide range of values is possible in a set of univariate traits does not imply that any combination of values among those traits are possible. In fact, an extensive body of literature in morphometrics and multivariate quantitative genetics shows that this is not the case (Klingeberg and Leamy, 2001; Polly, 2005). Morphological variation is often non-gradual, with the morphological structures limited to a number of discrete states (Jernvall, 2000; Szuma, 2002; Bronikowski et al., 2004). In addition, some variants occur more often than others, and some phenotypes may be very unlikely (or even impossible) to be generated by the development in a species in a given generation or even over a large number of generations (Jernvall, 2000). If this is the case, it follows that, in each generation, natural selection has a limited repertoire of forms to select from. Then the direction of evolutionary change depends on which variation is generated by development in each generation and which of it is picked by natural selection (Salazar-Ciudad, 2006).

Second, for natural selection to be the main explanatory cause in evolution, the relationship between genotype and phenotype (the so called Genotype-phenotype map or GPM) has to be simple. That is, genetically similar individuals have to always exhibit similar morphologies. If the relationship is not simple (complex GPM), then many phenotypes may be unreachable through selection (Marin-Riera and Salazar-Ciudad, 2013). This is because when the GPM is complex, many genotypes similar to the one producing the optimal phenotype may produce very different phenotypes. This causes that, in the adaptive landscape, there are many adaptive peaks (genotypes that produce high-fitness phenotypes) surrounded by adaptive valleys (adjacent genotypes that produce low-fitness phenotypes). Since natural selection only acts as a hill-climbing process in these adaptive landscapes, populations can only increase their fitness, going from adaptive valleys to adaptive peaks but never the contrary. That way populations can not cross from an adaptive peak

to another, higher (more optimal) one, thus remaining often trapped in suboptimal fitness peaks. Current evidence suggests that these complex genotype-phenotype maps are very common, being very likely to appear in many developmental systems (Salazar-Ciudad, 2001), and in many levels of the organism (Schuster et al., 1994).

Summarizing, despite the fact that mutation is effectively random, the phenotypes that arise from these combinations are not random (Oster and Alberch, 1981; Salazar-Ciudad, 2006).

All these facts pointed to a new reconsideration of the prominent role of development in evolution. This theoretical renewal did not happen until the '80s, contributing to start what it is currently known as Evo-Devo (Evolutionary developmental biology), a new sub-field of evolutionary biology which aims to understand how the organic forms (phenotypes) are generated by means of embryonic development and how they evolve. Furthermore, Evo-devo put again the organism as the fundamental unity of evolution (since it is the level in which natural selection can directly act) (Goodwin, 1996; Gould and Lloyd, 1999). Evo-devo has also an independent origin based on the wide application of molecular biology approaches to embryonic development (Gilbert, 2003; Jaeger et al., 2015). These two origins just happen to roughly coincide in time but are conceptually independent. Their integration is still taking place.

However, in its early form (often referred as “constraint school”), Evo-Devo viewed development as a limitation, constraint or burden imposed by the developmental dynamics to the total number of possible (attainable) phenotypes (Jacob, 1977; Gould and Lewontin, 1979; Alberch, 1982; Kauffman, 1993; Salazar-Ciudad, 2006). The hindering role of development in the form of the so called developmental constraint was summarized in the consensus definition: ‘‘A developmental constraint is a bias in the production of a variant phenotype or a limitation on phenotypic variation caused by the structure, characteristics, composition and dynamics of the developmental system’’ (Maynard-Smith et al., 1985). Notice that, at a deep logical level, natural selection continued to stand as the main evolutionary force, with development now precluding it to reach some of its potential phenotypical targets. Consequently, under this “constraint view” there were few insights on the mechanisms of the developmental processes (that is on how the combination of different developmental processes give rise to different morphological structures), mainly because researchers were more interested in what development does not allow to happen, rather in the things that happen during development (Salazar-Ciudad, 2006).

A more positive view of development as a generative force in evolution (and no merely as a “constraint” or “burden”) is also possible (Salazar-Ciudad, 2006). These authors claimed that, since the only way to create a given phenotype (at least for morphological phenotypes) is by means of development, then development does not impose a limit or constraint on evolution but, on the

contrary, allows it to happen through the production of phenotypical variation. This work stresses the fact that, for a developmental constraint to be an entity *per se*, one should then show how a constraint-free (that is, a development-free) phenotypic variation would be, which is actually impossible. Thus, the role of development is actually creative, and both development and natural selection jointly determine the direction of evolutionary change (the latter by carving in the phenotypic distributions produced by development). Thus, in order to understand how the phenotype of a lineage of organisms changes over generations, one must know not only which phenotypic variations are filtered out by environmental factors (natural selection); but also which phenotypic variation arises in each generation.

This constructive view about development allows us to decompose the complex process of development into a set of discrete stages and processes, so that one can understand how a stage gives rise to the next one, to compare different kinds of development between different organisms (establishing differences and commonalities), and to study the evolution of development itself. Some of these concepts, which will be relevant for further discussion on cleavage patterns and early development are briefly described in here.

Under this conceptual framework, development is the process whereby a single cell transforms, over time, into an organism composed of several types of cells arranged in a specific spatial pattern. This complex process can be conceptualized as a series of transformations of one developmental pattern (a specific spatial distribution of cell types in space), into other, usually more complex, ones. These pattern transformation usually involve genetic (gene networks) and cellular interactions; and any gene network that regulates at least one cell process (e.g. cell division, cell death, cell growth …) and it is involved in a pattern transformation is called “Developmental mechanism” (Salazar-Ciudad, 2003).

The gene networks (one gene or gene product affects (by up-regulating or down-regulating) the expression of others) often exhibit complex temporal and spatial dynamics capable on their own of pattern transformation. These developmental mechanisms which do not involve cell movements are called “inductive mechanisms”. On the other hand, it is often the case that a gene product (resulting from the gene network dynamics) regulates a specific cell processr that causes a change in the spatial positions of the cells. These developmental mechanisms are called “morphogenetic mechanisms” (Salazar-Ciudad, 2003).

Theoretical works carried out by Isaac Salazar-Ciudad (2003), Stuart Newman (Newman and Bhat, 2008), Jukka Jernvall (Jernvall, 2000), show that whichever morphological structure that organisms display can be built by means of a small set of developmental mechanisms adequately combined in space and time. Thus, any morphological difference between two individuals within a

species must arise first as a difference in some developmental mechanism at some developmental stage.

The discretization of development in a set of developmental mechanisms allows us to study two critical aspects of it. On the one hand the variational properties of such developmental mechanism, that is the range of possible phenotypes that it can be produce under different environmental perturbations and small mutations (those which do not alter the topology of the gene network of the mechanism considered, also called IS-mutations) (Salazar-Ciudad, 2006). On the other hand, the genotype-phenotype map (GPM) that is the relationship between specific genetic changes and specific phenotypes (Waddington, 1942).

However, the development of even rather simple organs often implies several developmental mechanisms, both inductive and morphogenetic, acting at the same time (morphodynamic), so that it is very difficult to disentangle the different processes involved, or to study the variational properties or the properties of the GPM for that organ. In these cases, computational models are useful in order to understand and analyze such complexity.

Computational modeling in Evo-devo.

Computational models of development have been extensively used to make the inherent complexity of development more manageable to the human mind. Ultimately, computational models are mathematical models that summarize the current knowledge of a specific developmental system, and that are computationally solved. In other words, a model requires a mechanistic hypothesis, posited in mathematical terms, of how an organ in a specific stage is built during development. When this mathematical model of the developmental system is computationally solved, the outcomes are usually the next developmental stage of that organ. Moreover, if one wants to raise one or several hypothesis on the developmental mechanism of that organ, he or she can change some of the parameters of the model, and see how the final phenotype looks like under these new mechanisms or mutations. This is said to be an *in silico* experiment.

In vivo analogs of these experiments can (and indeed are) carried out in living systems, but it is often the case that they are much more expensive and time-consuming. In other instances, e.g. if one wants to assess the whole range of phenotypes resulting from all possible IS-mutations in the system (that is, its variational properties), in vivo experiments are technically challenging, if not directly unreliable. Here is where the usefulness of computational models is manifested:

-They allow the researchers to identify the main factors involved in the generation of the wild-type phenotype (this is because, as in any model, a model of development is a simplification of the real system that only includes a subset of factors involved, isolated by physical and temporal boundaries). If the model successfully simulates the wild type phenotype, then is likely for those factors not-included in the model to be irrelevant in the development of the organ. In this case the model can be said to be validated, and one can not reject the mechanistic hypothesis implemented in it. Further experiments (in the lab), however, would be required to test or refine the hypothesis in the model.

-Compared to *in vivo* experiments, they allow a more integrative and systematic understanding of the developmental events because the parts of the model, as well as the physical, genetic and chemical phenomena involved can be easily isolated and quantified. The complex spatio-temporal dynamics of development becomes that way much more tractable, and subject to detailed analysis.

-By altering the parameters of the model (akin to mutation experiments in vivo), models allow to reproduce a vast amount of new phenotypic variants (this is it allows to know the variational properties of the system). By doing so intra or inter-specific phenotypic variation (as well as variation due to environmental changes) can be simulated, helping to identify the

developmental basis of the evolutionary transitions between them. If the model is again validated when confronted with real data, it may be able to predict the outcome of experiments that have not been yet carried out.

-However, if the outcomes of the model do not match those observed in in vivo experiments, the mechanistic hypothesis should be rejected and substituted by another one. In these cases, a rejection of the mechanistic hypothesis often provides a valuable insight about which are the aspects of the developmental dynamics that are more poorly understood and that need a reconsideration. This, in turn, may help to design future experiments or to modify the hypothesis on which the model is built.

-Computational models also allow to generate morphologies in a selection-free scenario. These developmentally possible morphologies, when spatially arranged according to some measurable features (e.g. Width, height, curvature ...), are called theoretical morphospaces (McGhee, 2001; Corominas-Murtra et al., 2013). Instead, when they are arranged according to the parameters that generate such final morphologies, they are called theoretical generative morphospaces (Niklas, 1999). The latter are much more informative since they show how the morphologies are distributed in the parameter space, thus it shows the structure of the genotype-phenotype map of the system. They are also informative because if there are some morphologies which are generated by the model (i.e. they are attainable through normal development) but they are not found in nature, it may mean that either there are a strong selective pressure that make these forms unadaptive (thus preventing their appearance in wild type populations); or there is some kind of historical contingency that makes these form very unlikely to appear (e.g. the current population(s) is drifting in a region of the parameter space that is very far away to those producing these phenotypes).

-When used along with in vivo experiments, computational models of development can be a very useful tools in assisting and guiding research programs, not only by providing new insights on how development works but also by saving money, resources and time.

Since the beginning of biological modeling, a large number of models have been developed, but not all of them have the same capabilities nor are they devoted to answer the same kind of questions related to development.

First models only implemented gene regulatory networks, without including cells or space as such (Thom, 1977; Kauffman and Levin, 1987; Kaneko, 1990). In addition, the gene concentrations, or the interaction strengths with other genes or gene products were often implemented in a binary fashion: either they are totally active (1) or inactive (0). The dynamics of

the GRN also relied on Boolean or trigonometric mathematical functions, that are able to reproduce to a certain extent some of the features found in biological systems. In these cases, the phenotype is the state (0 or 1) of each one of the genes involved in a given time. Despite their simplicity, these models provided interesting insights into the evolutionary dynamics of phenotypes exhibiting complex GPMs (rugged landscapes) (Kauffman and Levin, 1987).

Further models implemented the gene products concentrations and interaction strengths in a continuous manner, but again these models lacked the physical space in which the building of real phenotypes takes place (Wagner, 1994; Pinho et al., 2012). As in previous models, the arising phenotype is conceptualized as the levels of expression of genes, but in here it also arises explicitly from the non-lineal interaction between these genes, rather from *ad hoc* mathematical functions.

Other models, often considered to be developmental models, do not implement gene networks nor tissue dynamics per se, but instead try to link directly the genes with the phenotype (bypassing development as such). These models implement a genotype made of N genes and a phenotype made of P traits, and there are a number of lineal coefficients by which a gene *i* influences a trait *j*. Because of that this model is known as multi-linear model (Hansen and Wagner, 2001), and it was developed to simulate two well known phenomena: first, epistasis: a trait that is affected by various genes and two, pleiotropy; several traits that are affected by the same gene. Although it is true that the relationship between the genes and the phenotype (the genotype-phenotype map) is the main concern of development, the multi-linear model is strongly influenced by the field of multivariate quantitative genetics (MQG). The MQG approach establishes a statistical correlation between the observed phenotype and the underlying genes (thus being a up-bottom approach), without considering explicitly how development is acting in the molecular, cellular and tissue levels (black-box approach) (Lande and Arnold, 1983). Furthermore, the multi-linear model and some other related models (Gavrilets and de Jong, 1993; Nowak et al., 1997) assume that the number of traits is constant, and that genes have independent and linear effect on traits, that is in plain contrast with what is known about the developmental origins of phenotypic variation. Thus, these are not true developmental models and they are not going to be discussed further in this work.

Much more realistic and truly development-based models also exist. In these models, the phenotype arises from genetic interactions plus epigenetic factors and interactions between them. Epigenetic factors, *sensu lato*, are understood as anything that has a role in development but is not in the DNA sequence nor in the environment (Haig, 2004). Examples are asymmetric spatial distribution of proteins and RNAs in the oocytes of many species, gravity, the mechanical forces generated by tissues as they grow, etc. (Newman and Comper, 1990; Goodwin, 1994; Newman and

Muller, 2000).

Epigenetic factors are crucial in explaining the emergence of morphological structures in biology. Cells are physical entities capable of creating and responding to stimuli (they are excitable media), including the stimuli created by other cells. Because of that, when they coalesce in form of cell aggregates or tissues, the behavior of each cell is strongly influenced by the behavior of its cellular neighborhood. This interdependence makes that cell aggregates often exhibit a coherent collective behavior and self-organizing properties, which result in spontaneous pattern formation (Newman and Bhat, 2008). Ultimately, the factors that cause the cell aggregates to perform pattern formation (and to be more than a mere amorphous living matter) are epigenetic factors (e.g. diffusion rates of signaling molecules, cell adhesion, cell proliferation, forces exerted by the surrounding tissue …). Depending on the strength of each of these factors involved, the forms generated will be of one kind or another being some more likely to appear than others (Waddington, 1953; Alberch, 1982). Many of these forms, especially those that are more likely to appear (sometimes called “generic forms”), appear by means of simple physical processes and epigenetic factors, and do not require the direct control of genes to arise (Newman and Comper, 1990; Goodwin, 1996). The role of genes is then just to choose between the different easy-to-arise forms available by default, and the causative role in morphogenesis is displaced from naked genes towards the epigenetic factors or their interactions (Goodwin, 1996; Jablonka et al., 2007; Newman, 2012).

By implementing these epigenetic factors, models can lead to complex non-linear GPMs in which relatively small changes in the parameters can lead to relatively large phenotypic changes (Alberch, 1982). In these models, the phenotype is no longer the simple list of genes (or gene product) concentrations (as in the multilinear model), but a developmental pattern: a set of cells in particular spatial arrangement. Each of those may have a different state defined by their gene expression profile. Because of that, not all gene networks are capable of pattern formation (that is, capable of generating a phenotype), but only those whose internal dynamics, along with epigenetic factors considered, can change the cell states, the cell positions or both (Alberch, 1982; Salazar-Ciudad, 2006). Of those emergent phenotypes, some are quite likely to arise while others are not (Alberch, 1982), meaning that phenotypic variation has an intrinsic structure, what is in agreement with our current understanding of development (Alberch, 1982; Horder, 1989).

There are a large number of models including epigenetic factors, and they are applied to different biological levels; from RNA secondary structure (Schuster *et al.*, 1994) to neurophysiology (Skinner, 2012), but we are going to focus here in those applied to pattern formation and morphogenesis (Meinhardt, 1982; Honda *et al.*, 2008; Salazar-Ciudad and Jernvall, 2010; Osterfield *et al.*, 2013).

The simplest among these models is the so called lattice model, which includes an array of cells (each one containing a gene network), and the spatial diffusion of gene products between them (signalling) (Reinitz and Sharpe, 1995; Salazar-Ciudad et al., 2000, 2001; Jaeger et al., 2004). Thus, it can be considered as an extension of previously described Wagner's model in a spatial context (and thus including epigenetic factor of physical diffusion). As commented before, the main result of the analysis of the lattice model is that only a limited number of gene network topologies lead to pattern formation. In addition, networks with similar topology tend to produce similar morphologies (in other words, each type of network topology has an associated GPM) (Salazar-Ciudad et al., 2001).

However, development usually involves not only differences in gene expression but also differences in the location of cells themselves. In addition, both processes (cell signalling and cell movements) often occur simultaneously, in causal interdependence (Salazar-Ciudad *et al.*, 2003). Because of this, developmental models should implement morphogenetic processes (and the epigenetic factors involved). In order to achieve these cell movements, models should include the biomechanical properties of cells and some of the behaviors that they display during development: cell division, cell adhesion, cell contraction, apoptosis and changes in cell shape due to adhesion and contraction (Salazar-Ciudad et al., 2003).

A large number of developmental models do implement these things, but they can be classified in a few general groups according to the way cells are implemented. Each of these group has its own internal logic, advantages and limitations. Let us take a brief overview of the kinds of models:

The Cellular Potts Model (CPM) is a lattice-based model in which each cell occupies a number of adjacent cells in the lattice (Graner and Glazier 1992; Hogeweg 2000). Each part of a cell occupies a specific position in the lattice, and can move to another, adjacent lattice position by changing the identity of that specific position (the same holds when a part of a cell retracts from some position). This process allows the cell to change its shape and to move. Cell movements are computed based on some mathematical rules, which are commonly based on cell area (or volume in the case of 3D CPMs) conservation and cell-cell adhesion, so that the total energy of the system tends to be minimized. Different versions of the CPM include different cell processes, including cell-cell signaling, cell division, cell migration, apoptosis (Maree et al. 2007), but the model is not very suitable for those systems in which the cell shape, and the transmission of mechanical forces through the cells are relevant factors.

In the vertex model, cells are conceptualized as polyhedra (or polygons in the 2D case), and are defined by the contact interfaces with their neighboring cells (Honda et al. 2004, 2008). The

motion of cells are produced by changes in vertex positions, and cell rearrangements are performed by means of a vertex interchanges between adjacent cells. Vertex motions, in turn, are calculated by means of the Monte Carlo method or by a system of differential equations. As in CPM, the cell movement relies mainly on cell surface area/volume conservation, surface tension and cell-cell adhesion. The Vertex model has limited capabilities in modeling non densely packed cells or developmental processes where the cell neighborhood changes in a very dynamical way.

In the subcellular elements model (SEM) (Newman 2005, Sandersius and Newman 2008) each cell is composed of several subcellular elements (hereafter, nodes) that interact between them and with the nodes of the adjacent other cells. Each node has an equilibrium radius, a preferential distance it keeps with the other, adjacent nodes. Below or above this distance, repulsive and adhesive forces appear, respectively. The physical integrity of cells is ensured by keeping the intracellular attractive force between nodes larger than the intercellular one (Newman 2005).

As in the Vertex model, space is continuous, but in this model cells are defined by their volumes rather than by their contact interfaces with the adjacent cells, allowing for the implementation of a greater number of cell processes. Cell growth and cell death (apoptosis) can be readily implemented by changing the number of nodes of a given cell, and cell division by splitting the nodes of a cell into two daughter cells. Changes in cell shape can be produced actively (by modifying the equilibrium radius of the cell's nodes) or passively, by the rearrangement of the cell's nodes resulting from the basic forces of the model (Delile et al., 2013). Disadvantages include the difficulty of implementing epithelial cells and epithelio-mesenchymal interactions, that are known to be relevant in the development of many organs, such as teeth (Salazar-Ciudad, 2008; Salazar-Ciudad and Jernvall, 2010).

A new 3D modeling framework based on SEM, the EmbryoMaker, implements all elements (epithelia, mesenchyme and extracellular matrix (ECM)) and cell processes (cell adhesion, cell division, apoptosis, cell migration ...) that are known to be relevant in animal development (Salazar-Ciudad, 2003). In addition, these cell processes are controlled by gene products (diffusible or not) resulting from the dynamics of fully tunable gene networks. Both the initial conditions (the distribution of cells and ECM in space) and the mechanical properties of the cells (e.g. cell-cell adhesion) are also tunable. Thus, the EmbryoMaker is a truly general 3D model of development in the sense that it is not devoted to simulate any particular organ or developmental system. Instead, one can adjust the parameters of the EmbryoMaker to simulate any particular developmental system (this is because EmbryoMaker is actually a modeling Framework rather than a model per se).

To our knowledge, EmbryoMaker represents the state of the art in the computational models of development. It was developed by an international team of computational biologists under the

supervision of Isaac Salazar-Ciudad during the years 2011-2015, in both the Helsinki University and the Universitat Autonoma de Barcelona. The architecture of the EmbryoMaker is modular so that each involved researcher could develop independently a part of it. Most of the parts were programmed by Isaac Salazar-Ciudad and Miquel Marin-Riera, and my main contribution to it was concerned with the implementation of the biomechanics of mesenchymal cells, directed cell division, asymmetric mitosis and improvements in the graphical interface (all of them relevant parts in order to simulate cleavage, as it is explained below).

A case of study: the metazoan cleavage patterns.

So far, we have posited that Evo-devo challenges some of the current views about evolutionary biology, and that one valuable methodology in the Evo-devo research is the *in silico* experiment. However, one can only acquire a global understanding of how development works by getting new insights about the inner works of a specific developmental system (a developmental stage of an organ or structure). A specific developmental system, in which we will apply our computational approach, will provide us new data to be interpreted under the Evo-devo view about evolution. Ideally, this developmental system should meet some requirements. First, it should be a relatively simple system, involving a limited number of entities interacting during a short lapse of time. That way, it would be easier to disentangle the different factors (physical laws, epigenetic factors, cell processes, genetic regulation ...) involved in the system. Second, it should be widespread enough, and similar enough among several groups of organisms as to stablish comparisons between them. Third, it should be relatively well described at both mechanictic and descriptive levels as to implement the existing theories in a mathematical model.

In this Thesis, we focus on a specific developmental system that satisfies these requirements: the cleavage. Specifically, the goal of this thesis is to shed light on the developmental factors involved in the early cleavage of metazoans. Besides reviewing which are these factors in each animal group, we also have carried out a more detailed study for a representative type of cleavage pattern: the spiral cleavage pattern, by means of a computational model of cleavage. There are many reasons that explain the choice of this particular developmental system as a case of study, but before going deeper into these reasons, let us take an overview about the early development of metazoans: the cleavage stage.

Cleavage is the first stage of development in most animal species. It starts from a single cell, the oocyte (or egg), that despite being more or less spheric is far from being homogeneous. Normally, oocytes have internal asymmetries. Typically, some maternally inherited factors are present in one part of the cell (the vegetal pole) and not in the other (the animal pole). This asymmetry is described as an animal-vegetal axis. In many species another perpendicular axis, with different factors, also exists (Gilbert and Raunio, 1997). In addition, in many organisms the oocyte is surrounded by an outer cover or eggshell involved in protection, selective metabolite exchange and chemical attraction of the sperm cell. Once fertilized, a series of fast cell divisions partition the oocyte into a set of smaller cells called blastomeres. This is the process of cleavage, and in many species it proceeds without an overall growth of the embryo (the volume of the embryo is constant and roughly equal to that of the oocyte).

Yolk is a nutritive substance normally concentrated in the oocyte´s vegetal pole. Since yolk acts as an inhibitor of cell division, blastomeres in the vegetal part of the embryo divide more slowly than blastomeres in the animal part of the embryo (Gilbert and Raunio, 1997). Moreover, if yolk is dense enough it can not be pierced by the cleavage furrow when blastomeres divide, often resulting in incomplete cell divisions in which the cell nuclei and the cytoplasm of different blastomeres are not totally separated by cytoplasmic membranes. This type of cleavage is called meroblastic. According to the distribution of yolk within the blastula, meroblastic cleavage occurs in eggs that are either telolecithal (yolk is distributed throughout most of the blastula) or centrolecithal (yolk is located in the center of the blastula). Blastulae with a low or moderate amount of yolk display holoblastic cleavage (the furrows of cell divisions traverse the whole blastula, whereby blastomeres get individualized).

Cleavage finishes at the onset of gastrulation with a taxon-specific spatial distribution of blastomeres. The spatial arrangement of these blastomeres is what we call in here the "cleavage pattern". Many cleavage patterns exhibit geometrical regularities in their blastomere arrangement, which allows for a general classification into a small set of types according to the relative position of blastomeres (Gilbert and Raunio, 1997).

However, there is a weak relationship between cleavage patterns and phylogeny: distantly related taxa can exhibit similar cleavage patterns, while species belonging to the same taxon can have different cleavage types (Valentine, 1997). These discrepancies between the cleavage patterns and phylogeny have puzzled evolutionary and developmental biologists since Darwin times, and are difficult to explain and seem very counterintuitive unless we gain more knowledge on how the different blastomere arrangements can be generated during early development (Gilbert and Raunio, 1997). This means, how the spatio-temporal combination of different cell processes can generate the different cleavage patterns. This may provide valuable information about which cleavage patterns are more likely to appear by mutations affecting the cell processes involved in cleavage, and about the likelihood of evolutionary transitions between the different cleavage patterns (Alberch, 1982).

Thus, the main reason to choose the cleavage as a developmental system to study by means of computational modeling is because it is not well understood which cellular processes lead to cleavage. A computational model of cleavage will shed light on how different cell processes are combined during early development to generate the different types of cleavage patterns.

There are also other pragmatic reasons that make cleavage a very amenable system to be implemented in a computational model, namely: 1) Since during cleavage there is normally no cell-growth, the number of subcellular elements (the total number of spatial nodes where computations are taking place) in the model can be kept constant during simulations. 2) In most groups,

unambiguous taxon-specific cleavage patterns appear as soon as in the 16 or 32-cell stage, so that the number of cell divisions to consider when simulating the process of cleavage is rather low, shortening the time of simulations and saving computational resources. 3) During cleavage, the embryo is made of a single cell type: the blastomere, so one does not need to include in the model different cells (e.g. epithelial) or cells with different biomechanical properties (as we explain in the next section, we implement blastomeres as densely packed mesenchymal cells). 4) Blastomeres are relatively simple and undifferentiated cells, and they only display a limited repertoire of cell processes (see next section). Thus, the considered computational model can be simplified by including only these relevant cell processes.

A general realistic model of development, such as the EmbryoMaker (see previous section), would allow us to accurately implement the early cleavage taking into account all these considerations. This model is general enough to allow variations in the type and strength of the cellular processes involved, thus generating different cleavage patterns under different developmental parameters, and not only one particular cleavage pattern. By doing so, such a model may contribute to establish an integrative view of early cleavage, bringing different studies in the cleavage of different animal model systems under a common theoretical framework.

In order to implement the process of cleavage in a computational model of development, a detailed list of the main cell processes involved in metazoan cleavage, and how they are involved in the early development of the major metazoan groups is mandatory. Such a list, in the form of a review, constitutes the chapter 1 of this thesis work. The first part of this review offers an introduction to the generalities of cleavage, and the second part is a list of the cell processes known to be involved in metazoan cleavage. The third part of the review explains the main features of the cleavage patterns of all groups of metazoans, to the phylum level and, in many instances, to the class or order level. A summary of this review is to be published as a book chapter in: *Evolutionary Developmental Biology - A Reference Guide*, edited by Laura Nuño de la Rosa and Gerd B. Müller (Springer - Verlag). This review suggests that, despite the developmental diversity in cleavage patters, all metazoans seem to attain them by deploying the same basic cellular processes. Thus, the different cleavage patterns observed in metazoans (and the variation within each type of pattern) may be originated by mere variations in the kind and strength of these underlying cellular processes. Evolutionary transitions between different patterns, which have occurred many times in the evolution of metazoan lineages, would be driven by changes in these cellular processes (e.g. an adaptive increase in the amount of yolk, an inhibitor or cell division, can modify one the cleavage pattern until it lacks all the features of its original cleavage type, thus belonging now to another cleavage type).

Modeling the spiral cleavage pattern with SpiralMaker:

As commented in the previous section, a computational model of early development is very desirable, because it may shed light on how different cell processes are combined during early development to generate the different types of cleavage patterns.

To our knowledge, the number of computational models of early cleavage are scarce. Some of them focus on the bio-mechanical properties (rheology) of blastula-like aggregates and compared their results qualitatively with mammalian blastulae (Honda et al., 2008; Sandersius and Newman, 2008). Others (Bezem, 1975; Goodwin, 1980; Zammataro et al., 2007) are purely mathematical models based on geometrical (not developmental) rules, so that their biological insights they can provide are limited.

Finally, some of them try, as ours, to disentangle the morphogenetic processes responsible for the different types of cleavage existing in animals. The model that is most similar to the one presented in here is the work of Kajita and coworkers, but it applies only to one species (the nematode *Caenorhabditis elegans*), and it is restricted to the first two cell divisions (Kajita et al., 2003). In addition, it only includes cell division as basic cell process. Another recent model tries to explain the radial cleavage of the sea urchin (Akiyama et al., 2010). This latter model, however, is only 2D, and the number of cell processes it includes is not enough to simulate a large number of cleavage patterns beyond the sea urchin's one. Even for the case of sea urchin, this model only reproduces the 2D direction of cell division for some individual blastomeres at each cell stage, rather than the whole set of 3D directions of cell divisions within the blastula, as ours does. In addition, changes in cell shape due to surface tension and cell adhesion (that are crucial for establishing the direction of cell division) have to be manually introduced in this model for each cell stage.

In the chapters 2 and 3 of this thesis, we present a new model of early development that includes all cell processes that blastomeres display: the SpiralMaker. SpiralMaker is an application of the EmbryoMaker, the general model of development discussed in the previous section (Marin-Riera et al., 2015) to a particular developmental system: the early cleavage stages. For this application, some modifications have been carried out from the original EmbryoMaker:

-Simplifications: EmbryoMaker does not include epithelial cells, extracellular matrix nor the cell processes associated with these two elements, that are present in the original EmbryoMaker. Blastomeres are implemented as densely packed mesenchymal cells, and all cell processes other than cell adhesion, oriented cell division (by cell shape or intracellular gradients) and asymmetric

cell division have been deleted from the EmbryoMaker. In addition, the gene networks dynamics have been greatly simplified to mere internal gradients controlling some cell-behaviors, without gene-gene interactions as such. Because of these simplifications, SpiralMaker is no longer capable of simulating the development of different organs (as Embryomaker does), but only early cleavage stages.

-Additions. SpiralMaker includes some blastomere-specifc cell processes not implemented in the original EmbryoMaker. These are the so called Sachs' rule: an observational rule that accounts for the fact that each cell division tends to be perpendicular to the previous one (Minc and Piel, 2012); and the cortical rotation: a rotation of the blastomeres over themselves just after the cell division, which has been described in many developmental systems, and may have a role in changing the relative positioning of the blastomeres (Meshcheryakov and Beloussov, 1975; Wandelt and Nagy, 2004; Henley, 2012). A detailed explanation of how these rules have been implemented in the SpiralMaker are provided in the Annex 2 of this thesis.

However, the goal of using the SpiralMaker is not to simulate all known cleavage patterns, but only the spiral pattern (see Chapter 1 for a detailed description of this pattern). We decided to restrict our research to the spiral pattern by the following reasons:

-It is a conserved pattern: The spiralian cleavage pattern is the most abundant cleavage type at the phylum level. It is found in mollusks, annelids and nemerteans. Other Lophotrochozoan phyla (platyhelminthes, rotifers, brachiopods, phoronids, gastrotrichs, and bryozoans) also exhibit spiral cleavage in at least some of their species (Hejnol, 2010). In spite of having a very similar cleavage these phyla have very different adult morphologies. The ensemble of phyla with spiralian cleavage has been suggested to be a monophyletic group (Nielsen, 1994; Laumer et al., 2015), the *Spiralia*. In addition, some non-spiralian phyla also display a spiral-like (or pseudospiral) cleavage in some of their early stages (see chapter 1), so it is a widespread and representative type of metazoan cleavage.

-It is a well-known system, but its underlying biomechanics remains unclear: The Spiral pattern was already known for early embryologists, so it is a well known system with a considerable amount of available bibliographic data (Lillie, 1898; Wilson, 1898). Most of the literature related to the spiral cleavage, however, focuses in the signaling events that, taking place within a specific spatial blastomere arrangement, lay out the cell fate of each blastomere (Freeman and Lundelius, 1982; Kuroda et al., 2009; Grande and Patel, 2008). Much less is known about how the specific spatial arrangement of blastomeres in spiral cleavage is attained, and several (but a limited number

of) hypotheses have been proposed so far to explain how the spiral pattern is built by means of different cell processes. Some are roughly understood as developmental rules by which the direction of the cell division plane is determined during cleavage (Freeman and Lundelius, 1982) while others are processes of cell mechanical interaction that lead to cell displacements during cleavage (Meshcheryakov and Beloussov, 1975; Wandelt and Nagy, 2004; Henley, 2012). This allows us for include a limited number of well known cell processes, thus simplifying the model. By combining in the model these processes, and assessing which of them are capable, alone or in combination, of producing the spiral pattern, we can guess which of the previously held hypotheses contribute to explain the emergence of spiral pattern, and which do not.

-It exhibits limited morphological variation: Within spiralia, there exists considerable variation in the relative sizes and arrangement of blastomeres (all of them compatible with the spiral pattern) (Nielsen, 1994). All these different spiral patterns are thought to arise by small variations in some of the underlying developmental parameters (e.g. a grater amount of vegetally settled yolk can make the cells divisions more asymmetric, yielding bigger macromeres and smaller micromeres) (see Chapter 1). This morphological variation exhibited by the spiral cleavage pattern is enough for an initial survey of the variational properties of early cleavage without need to include all the known metazoan cleavage patterns. In order to explore this variation, and to assess how the different parameter combinations give rise to different spiralian cleavage patterns, we generate a (generative) theoretical developmental morphospace of possible spiral cleavage patterns. Each axis of this morphospace corresponds to one of the rules implemented in the model and along each such dimensions the relative contribution of each rule is quantitatively varied (see chapters 2 and 3 for details).

-Availability of real data: Within spiralia there is a handful of animal model systems species whose early cleavages have been described with a considerable accuracy. This means that relevant data concerning the relative positioning or the relative sizes between blastomeres are readily accessible. Given that the outputs of the simulations of the model we use -the SpiralMaker- are realistic 3D representations of a 16 (and 32) cell-stage embryos, we can compare our simulated blastulae with those of real spiralians. In order to do so, we have developed two different methods (one based on the relative volumes between blastomeres and the other on the relative contacts between adjacent blastomeres). By using these methodologies (see chapters 2 and 3), simulated cleavage patterns reminiscent of those of several spiralian species (*Crepidula*, *Planorbella*, *Lottia*, *Trochus*, *Carinoma*, *Nereis* and *Arenicola*) were found within our theoretical developmental morphospace.

-It is enough for a first approach: As commented before, the number of computational

models of early cleavage is rather low, and within them, the ones capable of generating inter-specifc variation in these cleavage patterns are, to our knowledge, virtually inexistent. Thus, for the sake of simplicity, it may be better to restrict this first approach to the widespread, representative and relatively well studied spiral pattern.

The modeling of the spiral cleavage pattern with the Spiralmaker has provided two main insights (further explained in the chapter 2 of this Thesis):

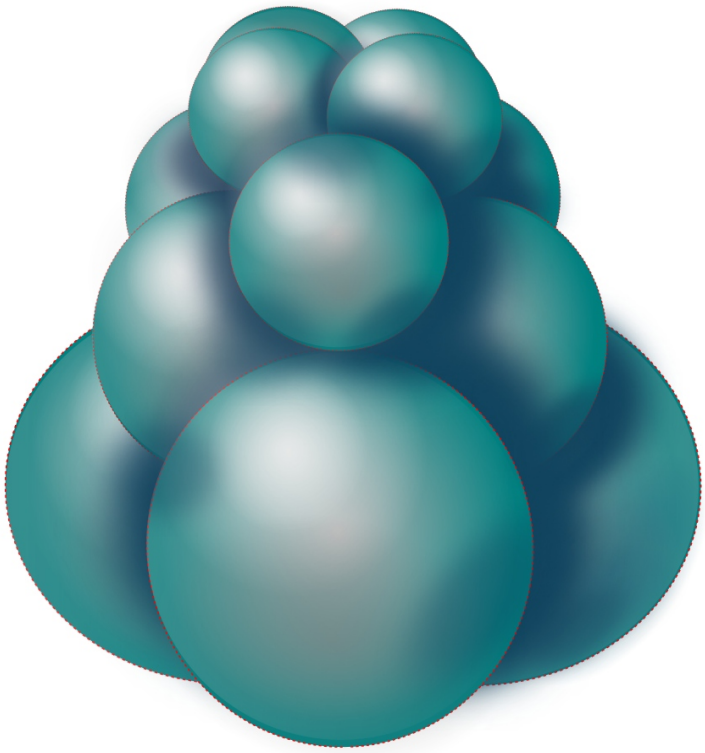
First, when the different cell processes are implemented and combined in SpiralMaker, most of the combinations do nor lead to the spiral pattern. The only combination in which the spiral pattern arises is the one that includes the cell polarization rule (oriented cell division along the Animal-Vegetal axis), the Sachs' rule, cortical rotation and inter-cellular adhesion. Specifically, cortical rotation is relevant in the third division to produce a coherent twist or rotation of all micromeres respect to the macromeres, and Sachs' rule is responsible of the left-righ alternation of cell divisions after the 8-cell stage (see chapter 2).

Second, our work shows that part of the morphological variation observed in spiralian cleavage can be explained by variations in the parameter values of these few rules. Species-specific patterns, reminiscent of those of seven real species were found in the morphospace by means of different procedures. This suggests that the different spiralian patterns arise for particular parameter combinations, being some of them more likely to arise than others. Thus, in spite of its simplicity, our model seems to be indicative about the developmental processes underlying morphological differences at the level of blastulae between species.

CHAPTER I

The evolution of cleavage patterns ine metazoans

A summary of the review which constitutes this chapter is to be published as a book chapter in: *Evolutionary Developmental Biology* - A Reference Guide, edited by Laura Nuño de la Rosa and Gerd B. Müller (Springer – Verlag). (Submitted February 2016)



THE EVOLUTION OF CLEAVAGE IN METAZOANS

Miguel Brun-Usan[a], Isaac Salazar-Ciudad[a,b,*]

[a] Genomics, Bioinformatics and Evolution. Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Barcelona, Spain.

[b] Evo-devo Helsinki community, Center of excellence in computational and experimental developmental biology, Institute of Biotechnology, University of Helsinki, PO Box 56, FIN-00014 Helsinki, Finland.

* Corresponding author: isaac.salazar@helsinki.fi

Abstract

Cleavage is the earliest developmental stage. In it, the fertilized oocyte gives rise to a cluster of smaller cells (blastomeres) with a particular spatial pattern (a cleavage pattern). Different metazoan species have different cleavage patterns, but all of them fit into a small set of basic types.
The relationship between the phylogeny of a given species and its cleavage pattern is far from direct, but most taxa seem to use the same basic cell processes (such as directed cell division or cell adhesion) to build their cleavage patterns. In first section of this chapter, we assess which are those mechanisms.
In a second section, we explore how the combined action of these mechanisms can account for the emergence of particular cleavage patterns in different metazoan taxa and the evolutionary transitions between them.

Keywords: early development, cleavage pattern, cell processes

1. Introduction.

Metazoans display a wide range of reproductive strategies (such as gemation in acoel flatworms, viviparism in placental mammals and parthenogenesis in eusocial insects). But, most of them, in at least one stage of their life cycles, pass through a single-cell stage. This single-cell stage (egg or oocyte) contains all the information to, given the appropriate environmental conditions, build a functional adult organism. If a given species undergo sexual reproduction, the oocyte has to be fertilized by another cell (the spermatozoa) in order for development to start (and then the oocyte is called zygote), but in parthenogenetic species (a mode of asexual reproduction that does not involve “fathers” or spertmatozoans) it can proceed by itself. In any case, the building of an organism from a single cell stage is produced by means of a complex process called development (Gilbert and Raunio, 1997).

Basically, development can be conceptualized as the transformation of one developmental pattern (a 3D distribution of cell types in space) into another (Salazar-Ciudad et al., 2003). These transformations are mediated by developmental mechanisms. We define developmental mechanism as a gene network capable of pattern transformation, which usually involves the regulation of at least one cell behavior (e.g. directed mitosis, cellular adhesion ...). The number of known developmental mechanisms (e.g. directed mitosis, apoptosis, cellular adhesion) is relatively low.

Whichever morphological structure observed in animals, irrespectively of its complexity, is generated by a specific spatio-temporal combination of these few developmental mechanisms. Developmental mechanisms are also active in the first developmental stage, that is called cleavage.

Cleavage starts from a single cell, the oocyte (or simply egg), and finishes at the onset of gastrulation (the next developmental stage, when the multi-layered nature of the adult organisms is established). The oocyte, despite being more or less spheric is far from being homogeneous. Oocyte ´s inhomogeneities normally consist on uneven (and non random) distribution of gene products within the cell cytoplasm (Raven, 1967). These products are often internal gradients (maternally inherited), that are oriented along one axis called animal-vegetal axis and sometimes also in a dorso-ventral axis (Gilbert and Raunio, 1997)). The "vegetal" hemisphere of the embryo is termed that way because of its apparent immobility respect to the animal half, where the cytoskeletal dynamic is more active during cleavage (animal pole is normally enriched with subcortical actin). The nucleus of the oocyte is commonly located towards the animal pole. Moreover. In many embryos, the animal pole is marked prior to fertilization by two tiny haploid peripheral cells (polar bodies) that are also produced during oogenesis.

During cleavage, a series of fast cell divisions partition the oocyte into a set of smaller cells called blastomeres, and proceeds without an overall growth of the embryo (the volume of the embryo is constant and roughly equal to that of the oocyte). In many organisms, the oocyte is surrounded by an a more or less rigid coverage and then part of early development occurs within the space enclosed within this coverage. It ranks from the soft hyaline layer of mollusks and sea urchins to the rigid eggshells of nematode worms. The functions of these covering structures include protection, selective metabolite exchange and chemical attraction and reception of the sperm cell (e.g. the zona pellucida in mammalian eggs or the egg jelly in sea urchins, (Gilbert and Raunio, 1997)). The possible role of this coverage concerning the arrangement of blastomeres will be discussed in the section 1.3.2. (In this work, the term eggshell refers this coverage, irrespective of its stiffness or hardness, as it is customarily used in the literature).

During cleavage, the embryo is termed blastula (or, in less often, "morula" because of its resemblance to a morus, mulberry). The spatial distribution of blastomeres within the blastula is what we call a "cleavage pattern". Different metazoans species show different cleavage patterns. The cleavage patterns found in metazoan can be classified into a small set of types according to some general features, such as the amount and distribution of yolk or the relative position of blastomeres (Gilbert and Raunio, 1997).

Yolk is a nutritive substance normally concentrated in the oocyte´s vegetal pole. At evolutionary level, it is a derived feature (holoblastic eggs were ancestral) devoted to nourish the

developing embryo when it does not feed by itself (Siewig, 1979). Since yolk acts as an inhibitor of cell division, blastomeres in the vegetal part of the embryo divide more slowly than blastomeres in the animal part of the embryo. Moreover, if yolk is dense enough it can not be pierced by the cleavage furrow when blastomeres divide, often resulting in incomplete cell divisions in which blastomeres are not totally separated by the cytoplasmic membrane. This type of cleavage is called meroblastic. According to the distribution of yolk within the blastula, meroblastic cleavage occurs in eggs that are either telolecithal (yolk is distributed throughout most of the blastula) or centrolecithal (yolk is located in the center of the blastula). Blastulae with a low or moderate amount of yolk display holoblastic cleavage (the furrows of cell divisions traverse the whole blastula, whereby blastomeres get individualized). In some case, however, the correlation between the amount of yolk and the completeness of the cell division is not so straightforward: there are examples in which heritage overrules the putative mechanistic effects of yolk mass (Fioroni, 1987), and other in which very yolky eggs undergo holoblastic cleavage (Scholtz and Wolff, 2002).

In addition, many cleavage patterns exhibit geometrical regularities in their blastomere arrangement, which allows for a general classification into a few major cleavage types. These regularities are more patent when the limits between blastomeres are well defined (i.e. in holoblastic cleavage), One of the main regularities that are used for cleavage types classification is the angle between pairs of sister blastomeres and the main (animal-vegetal) axis of the blastula. That way, cleavage patterns are usually classified as:

i) Radial cleavage pattern (or radial-like): the directions of the successive cell divisions are always parallel or at right angles to the animal-vegetal axis. Along this axis, blastomeres are located one on the top of each other.

ii) Spiral cleavage pattern: the directions of cell divisions are oblique respect to the animal-vegetal axis after the second cell division.

Both spiral and radial cleavages proceed in a similar manner up to the 4-cell stage: as in other types of cleavage, they begin with two successive nearly meridional cell divisions giving rise to four large blastomeres that lie in a plane perpendicular to the animal-vegetal axis.

In other instances, the symmetry axes showed by the blastulae are used for their categorization. Actually, the cleavage (and development itself) can be seen as a set of symmetry-breaking events that transform a spherically symmetrical egg into a multicellular architecture with

fewer symmetry axes (Zarenkov, 2005; Presnov et al., 2010; Isaeva et al., 2012). In some cases, the symmetry axes of the early blastula correlates with the adult organism´s body axes. According to their main planes of symmetry, we can distinguish between:

i) Bilateral cleavage pattern: The first cell division splits the egg into left and right halves, creating a plane of bilateral symmetry. The following cell divisions are symmetric respect to this plane, and cell divisions in the right half are mirror images of these of the left half. This results in the two halves of the blastula being mirror images of one another.

ii) Biradial cleavage pattern: As in bilateral, but involving an extra plane of symmetry perpendicular to the primary one.

iii) Rotational cleavage pattern: In this case, after the first cell division creates a plane of symmetry, each cell divides perpendicularly to the previous cell division (parallel to the plane) and perpendicular to each other. That is: one cell divides meridionally and the other equatorially. In rotational cleavage, both pairs of blastomeres only look the same after certain amount of rotation around a rotation axis that is normal to the plane of the first cell division.

These few categories do not cover all the diversity, but represent major themes, corresponding to the more specious and more studied phyla. These major themes admit many exceptions and some degree of variation within. Other minoritary, more specific types of cleavage, characteristic of smaller and less studied taxa are also considered here, but will be described in the second part of this work. It worths also mention that if two different species share the same type of cleavage, it does not imply that the underlying developmental mechanisms, the mode of cell fate specification or the cell fate map were the same between these species (this evolutionary substitution of one developmental mechanisms by another while keeping the same final morphology is often referred as “developmental system drift” (or DSD) ((True and Haag, 2011), see for instance the different cell processes that lead to an asymmetric cell division in Section 2.1.3). Thus, a given type of cleavage pattern can not be used as a single “character” state for phylogenetic inferences, but is a complex trait itself should be decomposed in many finer characters (homologies between different blastomeres, mechanisms involved …) in order to compare between different patterns (Jenner, 2003; Jenner, 2004; Donoghe, 2005).

Distantly related taxa can exhibit similar cleavage patterns (e.g. both mammals and nematodes have rotational cleavage), while species belonging to the same taxon can have different

cleavage types (e.g. Gastropod mollusks such snails undergo holoblastic spiral cleavage, but their close relatives the cephalopod mollusks have meroblastic telolecithal cleavage) (Gilbert and Raunio, 1997). This results in a weak relationship between cleavage patterns and phylogeny. (Notwithstanding this, at coarse level it is possible to arrange the phyla to produce a parsimonious distribution of cleavage types (Valentine, 1997), being the radial cleavage the ancestral condition, but such a simply schema is fraught with exceptions and ambiguities).

This lack of direct phylogenetic correspondence between cleavage patterns is related to a phenomenological developmental model known as the “hourglass model” of development. The logic of the hourglass model can be summarized as follows: first, different cleavage patterns often give rise to almost identical morphologies in the intermediate developmental stages. Second, after this morphologically conserved stage (also called phylotypic stage), the embryo morphologies of the different species of the taxa get again more and more different as development proceeds. In other words, development diverges more extensively early and late than in the middle, so that if we plot in the Y coordinate the developmental time and in the X coordinate the morphological variability we obtain a profile that resembles an hourglass: wide in the basis and in the top, and narrow in the middle. The observed variety in cleavage patterns among metazoans would then correspond to the wide basis/lower part of the "hourglass". The extension and the evolutionary significance of the hourglass model is still discussed see (Domazet-Lošo and Tautz, 2010; Newman, 2011; Levin et al., 2016). While many efforts have been made to disentangle the evolutionary forces that constrain the phylotypic stage, the forces that produces the initial variability in the lower part of the hourglass (that is, the different types of cleavage patterns) are less studied.

Thus, the variability of the cleavage pattern itself, and the discrepancies between the cleavage patterns and the phylogenetic position of metazoan taxa are difficult to explain and seem very counterintuitive unless we gain more knowledge on how the different blastomere arrangements can be generated during early development. This means, how the spatio-temporal combination of different cell processes can generate the different cleavage patterns. This may provide valuable information about which cleavage patterns are more likely to appear by mutations affecting the cell processes involved in cleavage, and about the likelihood of evolutionary transitions between the different cleavage patterns (Alberch, 1982; Valentine, 1997). The aim of this work is to gain an overview of the main cell processes involved in metazoan cleavage, and to review how they are combined in the early development of the major metazoan groups. Few works covering the early development of all metazoan clades have been carried out, and none of them (to our knowledge) considers explicitly the biomechanical process involved in cleavage. Most of them are either encyclopedic treatises (Grassé, 1975) or books devoted only to the main model species (Strathman,

1987; Gilbert and Raunio, 1997; Müller et al., 2015) rather than comprehensive reviews. In others, the different chapters (devoted to different taxa) have been written by different authors, precluding a common and equitative approach to all taxa (Wanninger, 2015).

2. Cell processes involved in cleavage.

Compared to later developmental stages, during cleavage only a single kind of relatively undifferentiated cells exists, and the number of cell processes that they can display is relatively small. As explained above, many authors have argued that all diversity of morphologies found in metazoans results from the spatio-temporal combination of a small repertoire of very basic cell processes (Salazar-Ciudad, 2003; Newman and Comper, 1990; Newman, 2012). This also holds for the cleavage patterns, given the fact that the number of cell processes that blastomeres can display is rather low. These cell processes are:

2.1. Cell division.

Cell division can occur in a specific direction (the plane of cell division can be oriented in different ways in space), specific moment in time and with a specific degree of size asymmetry between daughter cells (Gillies and Cabernand, 2011).

2.1.1. The direction of cell division:

In this work, we use the term “direction of cell division” as the direction of the imaginary line that links the two centrioles lying in opposite sides of a cell during cell division, each one in one extreme of the mitotic spindle. This orientation of the mitotic spindle and the associated cytoskeletal machinery involved in cell division is known to correlate with the formation of the contractile ring, and ultimately with the relative spatial position of the two resulting daughter cells (McNally, 2013). Thus, an alternative definition of the direction of cell division would be the hypothetic line that will link the centroids (center of masses) of the two resulting daughter cells. Less often we use the term “plane of cell division”, that is the imaginary plane that separates the daughter cells just after cell division has been completed. By simple geometry, the plane of cell division is always perpendicular to the plane of cell division.

Cells can be polarized and this polarization can determine the direction in which cells divide. In order to be polarized, cells require one or several sources of spatial information (actually,

three bits of spatial information: x, y and z coordinates are enough to specify a direction in a three-dimensional space). Ultimately, cell polarization results when one or several sources of spatial information are translated into a spatially asymmetric distribution of some specific molecules within the cell (Freeman and Lundelius, 1982; Lu and Johnston, 2013). In many cells under division this asymmetry promotes a differential attachment of the astral microtubules to the part of the cortex with the highest concentration of these molecules, thus tilting the mitotic apparatus and biasing cell divisions to occur along the direction of cell polarization (Freeman and Lundelius, 1982; Lu and Johnston, 2013). The attachment of the astral microtubules to the cortex is not direct, but is mediated by other factors such as dynein (a motor protein) that are responsible of the pulling force on microtubules (McNally, 2013). As explained before, many molecular species, including some with polarizing effects, are aligned along the animal-vegetal axis of the blastula. Because of this, many cell divisions are oriented either along the animal axis (equatorial cell division) or perpendicular to it (meridional cell division). The first cell division is meridional (the plane of cell division bisects both animal and vegetal poles of the egg) in most of metazoans, irrespective of their cleavage pattern.

The sources of spatial information can be located either within the cell itself (autonomous mechanism) or in the cell´s surroundings (inductive mechanisms). The autonomous mechanisms do not require any physical or chemical interaction with other cell(s) to determine its direction of cell division. That means that cells use asymmetries that are already present in the intracellular environment to polarize themselves. These asymmetries are usually inherited, and usually consist in the heterogeneous distribution of some factor(s) in the cytoplasm. In other cases these heterogeneity is attained when some factors (if dense enough) are attracted by gravity to the lower part of the embryo, creating new spatial information. These factors can be diffusible molecules (thus forming an intracellular gradient) or non diffusible aggregates (thus forming dense granules). In this latter case, if factors are only present in a very restricted subvolume of the cell, they can eventually be inherited by only one daughter cell. These factors can interact (directly o by means of other intermediate factors) with the mitotic spindle in such a way that the spindle always tends to point towards the place where the factors are most abundant (Day and Lawrence, 2000).

When other spatial cues are absent, the cell shape itself (specifically, its longest axis) is able to determine the direction of cell division (Besson and Dumais, 2011). This is commonly referred as Hertwig's rule, and it has been described in many developmental systems (specially in proliferating epithelia). Hertwig´s rule occurs because tension in astral microtubules depends on the contact angle between the microtubule tips and the cell's surface. By simple geometry, this angle is smallest at the cell boundary at the most distant points of the cell, causing that the astral microtubules

attached there exert a stronger tension than the microtubules attached elsewhere, thus leading to the alignment of the mitotic spindle along the longest axis of the cell (and then to cell division to occur along that axis) (Théry and Bornens, 2006; Gillies and Cabernand, 2011; Minc and Piel, 2012, Besson and Dumais, 2011). Concerning cleavage, it has been proposed (Meshcheryakov, 1978) that, in some taxa, Hertwig's rule could translate blastomere's shape changes (for instance due to the packing produced by cell-cell adhesion) into specific directions of the cell division planes.

Finally, there exists a phenomenological rule (known as Sachs' rule) by which the direction in which a cell divides tends to be perpendicular to the direction of the cell division that gave rise to it (that is its mother cell division). This has been proposed to arise from the stereotypic (90º) duplication of the centrioles between cell divisions, that in turn biases the position of the mitotic spindle towards perpendicularity (Théry and Bornens, 2006; Minc and Piel, 2012). However, when directions of cell divisions are not restricted to take place within a plane (such as in epithelia), Sachs' rule alone can not unequivocally determine a direction of cell division. This is because in 3D space there are an infinite number of directions (albeit in the same plane) that are perpendicular to the direction of the previous division). In that sense, if Sachs' rule holds in cleaving embryos (as it has been proposed for some animal groups Guerrier, 1970), it can not specify the direction of cell division unless it was “combined” with another source of spatial information.

Other way for cells to choose a unique direction of cell division among the many options provided by the Sachs´ rule is the so called Errera´s rule (Errera, 1882). Errera´s rule (another classical phenomenological rule) posited that cells, when setting a new boundary, minimize their surface area and the predicted division plane among those that respect Sachs’ rules is thus the shortest possible one (Minc and Piel, 2012). However, Errera´s rule predictions are oftentimes challenging to test in non-trivial cell geometries, such as in blastomeres. This is probably the cause the reason because it has only only been proposed to be relevant in vegetal cells, with a clear polyhedral geometry (Besson and Dumais, 2012).

In the case of inductive (non-autonomous) mechanisms, the direction of cell polarization is affected by clues in the cell's surroundings. That is, the dividing cell extract the spatial information from outside. In many cases, this is achieved by means of short-range (Lu and Johnston, 2013) or long-range (Green et al. 2008) diffusible signals between neighboring cells (inductive mechanisms). A special case of “short range” signals is the physical contact between cells. The direction of cell division is also known to be affected by which parts of a blastomere are in contact with other blastomeres In this case, cells tend to divide towards (or against in some cases (Wang et al., 1997)) the part of the cell making contact to adjacent cells. This has been suggested to occur because

physical contact in a cell region would modify the underlying cell cortex so that the astral microtubules are stabilized in this region, increasing the local traction of the mitotic spindle during cell division (Théry and Bornens, 2006; Goldstein, 1995; Hertzler et al. 1992). Polarizing effects caused by factors not belonging to the blastula itself has also been documented. For instance, nurse cells found in many insects or the ovary cells of *C. elegans* are capable of printing an anteroposterior polarity to the developing embryo (Goldstein, 2001; Chipman, 2015). The physical contact with a physical substratum may also be used as a polarizing mechanism, (Watanabe, 1978).
Finally, another external source capable of orienting cell division (albeit only the first one) is the entry point of the sperm cell into the oocyte. When it occurs, the sperm cell pronucleus promotes a cytoplasmic flow within the oocyte that redistributes some cytoplasmic factors (e.g. PAR proteins in *C. elegans*) in an asymmetric way. This asymmetry ultimately polarizes the cell and guides the cell division.

2.1.2. Differential growth:

During cleavage cells in a blastula can divide at the same time (synchronous cell divisions) or not (asynchronous cell divisions). Synchronous cell divisions give rise to different cleavage patterns than asynchronous cell divisions. In general the resulting cleavage patterns depend on the relative rates of cell divisions between the different regions of the blastula. Assuming that some factors can trigger (or inhibit) cell division, this asynchrony can be achieved by an heterogeneous distribution of those factors in different regions of the blastula.

One of these factors is yolk (which usually forms an animal-vegetal gradient), that is known to delay (or even inhibit) cell division. As a consequence in many cases cells close to the vegetal pole divide at a slower pace, and remain bigger than those close to the animal pole (notice that this difference in the relative size between animal and vegetal blastomeres is not due to asymmetric mitosis). In some groups, cell division is inhibited just in a single specific blastomere which becomes larger than the others.
By default, the timing of cell division rounds in cleavage is regulated by a set of a family of cytoplasmatic proteins called cyclins (synthesized by maternally provided mRNA) (Evans et al., 1997). Cyclin expression follows an oscillatory temporal dynamics coupled with cell cycle. This oscillatory dynamics can result from closed chains of molecular events that trigger each other in a sequential fashion or by genetic networks with inherent oscillatory dynamics (molecular clocks) (Murray and Kirschner, 1989).

2.1.3. Specification of daughter cells' size:

In general, when a cell divides the resulting daughter cells are equally sized (symmetric cell division), but in many embryos some cell divisions are asymmetric: one daughter cell (macromere) is significantly larger than the other (micromere). When there is in addition cells of intermediate size interspread they are called mesomeres. Different mechanisms can result in asymmetric cell division. In some cases, the relative size of daughter cells is regulated by the asymmetric concentration of intra-cellular factors (Gilbert and Raunio, 1997; Salazar-Ciudad et al., 2003; Munro et al., 2006). These factors (e.g. PAR proteins in *C. elegans* embryos (Cowan and Hyman, 2004)) polarize the cell cortex, making that its properties are no longer equal along the cell (McNally, 2013; Ren and Weissblat, 2006). Then, the microtubules of the mitotic spindle get more stabilized in the regions of the cell cortex where these factors are more abundant, generating asymmetric pulling forces during cytokinesis. Since the concentration of these factors often varies along the animal-vegetal axis, cell size gradually increases along the animal-vegetal axis (Actually, the correlation between the cells´ size and their relative position along the animal-vegetal axis is so strong that some researchers considerer micromeres as the cells being closer to the animal pole (macromeres to the vegetal), irrespective of their relative size (Hejnol, 2010). Aligned or not along the animal-vegetal axis, these factors involved in asymmetric cell division interact with the spindle microtubules, generating asymmetric pulling forces and a displacement of the contractile ring during cytokinesis.

Alternatively, asymmetric cell division can result from an inherently asymmetric spindle. In these cases (e.g. *Tubifex* worms (Ren and Weissblat, 2006)), one centrosome is inactivated, so as only a half of the mitotic spindle is plenty developed and exert grater traction force than the other, degenerated, half of the spindle, thus displacing the cleavage plane to one side of the cell.

Asymmetric mitosis may play a role in generating variation between cleavage patterns (Merkel et al., 2012). This is because when the cells are tightly packed (by cell adhesion and/or by an eggshell limiting the available space for new cells to fit), the resulting blastomere arrangement may depend on the relative size between blastomeres (e.g. small blastomeres may occupy the furrows between big ones). This increases the number of possible spatial architectures of blastulae.
Cell division is also said to be asymmetric if the mother cell has some kind of internal polarity (e.g. an mRNA gradient), and the two daughter cells inherit it in a differential manner, even if they are equally sized. That way one daughter cell can incorporate different molecules than the other, which may cause differential gene expression between sister cells.

2.2. Cell processes not related to cell division:

These mechanisms does not specify the direction of cell division per se, but it may affect the mechanical interactions between cells after division and lead to changes in the relative positioning of the blastomeres.

2.2.1. Cell adhesion:

During cleavage, cell adhesion keeps the blastomeres together, thus maintaining the physical integrity of the blastula as a whole. Moreover, cell adhesion increases the contact surface between adjacent cells, which can lead to cell shape changes that may affect the relative position and contacts between neighboring blastomeres (Lecuit and Lenne, 2007), and even the direction of the cell divisions if Hertwig's rule applies. If adhesion molecules are expressed non-uniformly on the surfaces of individual blastomeres, complex spatial arrangements, such as embryonic cavities (lumens) or cell chains can be formed (Newman and Tomasek, 1996; Shinbrot, 2009). In some taxa, adhesion strength is not constant over cleavage time, but cells suffer cycles of increased cell adhesion coupled to cell division cycles (Bezem and Raven, 1975; Fritzenwanker et al., 2007).

Local variations in cell adhesion (and in the surface of contact between blastomeres) are important when the cell fate determination is controlled by inductive mechanisms. In these cases, cells are not induced below a certain area of contact, but are only induced above this area (Munro et al., 2006).

2.2.2. Cortical rotation:

During the first cell divisions in certain taxa, blastomeres rotate over themselves just after cell division around the rotation axis that links the two cells (Meshcheryakov and Beloussov, 1975; Wandelt and Nagy, 2004; Henley, 2012). Around this axis, rotation occurs in the same sense in all blastomeres (e.g. all counterclockwise respect to their sister blastomere, as if they were “unscrewing”). Whereas in some taxa this rotation does not seem to have any morphogenetic effect (Danilchick et al., 2006), it has been suggested that this rotation produces relevant changes in cell relative positions in other taxa (see section 3.4) and the sorting of cytoplasmatic determinants that are essential for further development (Gerhat et al., 1989). The molecular mechanics of this rotation remains unclear but F-actin seems to be a major player (Danilchik et al., 2006). F-actin molecules have chiral structure which may produce a preferential twist in the traction forces exerted by the

cytoskeleton on the cortex (Henley, 2012; Tee et al., 2015). Moreover, F-actin exhibits self-organization properties that favors the coordinate assembly of its filaments in the same direction, rather than the formation of randomly oriented bundles (Tee et al., 2015). Bot chirality and self-organization have been proposed to create a directional rotational flow of actin in the cell cortex while the contractile ring is stretched during cytokinesis, causing the cortical rotation (Hird, 1993; Danilchik et al., 2006; Tee et al., 2015).

2.2.3. Packing constraints:

As discussed above, it is often the case that cleavage proceeds inside a more or less rigid eggshell. When cleavage proceeds inside an eggshell, a compressive effect may be exerted by the limitation of the available physical space for the blastomeres (Schierenberg and Junkersdorf, 1992; Kajita et al. 2003). Due to geometrical considerations (Zammataro et al., 2007), when a set of spheres (blastomeres) is packed within a limited three dimensional space, there are only a small number of optimal cell spatial arrangements. These optimal arrangements are minimal-energy configurations, and depend on the shape of the eggshell. Thus, packing constraints could facilitate the constant positioning of the micromeres within the blastulae (Wang et al., 1997; Goldstein, 2001; Kajita et al., 2003; Zammataro et al., 2007, Isaeva et al., 2012).

Moreover, if the eggshell itself is not spheric (as it is the case in *C. elegans, Drosophila*...) or it has some structural asymmetry it readily provides a source (one bit) of spatial information maternally inherited, which can be used for the mechanisms described in the section 1.3.1.2.

Notice that none of the previously described mechanisms is directly encoded genetically (Newman, 2011). Rather, they arise from epigenetic factors (*sensu* Newman and Muller, 2000). Under this interpretation, epigenetic factors are not only restricted to those elements that are not DNA but regulates its transcription (such as histone modification or DNA methylation), but include other non-genetic elements that are causally involved in the generation of biological forms. This wider definition of epigenetic factors incorporates the complex dynamics of the cell cytoskeleton and purely physical processes like membrane surface tension, volume displacement, gravity and molecular diffusion. It is worth mentioning that some of these processes also apply to inanimate matter. Because of that, some non-living systems such as soap-bubbles or mineral aggregates share many geometric regularities with cleaving embryos (Thompson, 1917). For instance, Errera (1886) stated that the boundary of a cell adopts the geometry that a soap film would take under the same conditions (Besson and Dumais, 2012). The comparison between real cleaving embryos and those

blastula-like non-living systems is not meaningless because, in the later, not evolutionary history nor adaptation does exist, and the mechanisms involved are just simple mechanical laws. The “cleavage patterns” in those inorganic systems inform, thus, of which are these “default” cleavage patterns that can arise by means of mere physical principles, without additional genetic control. These cleavage patterns, by requiring less precise regulation, are more likely to arise in evolution (since they require less mutational changes).

Moreover, most of the cellular processes described in here, as well as their molecular bases are not specific of blastomeres, but are also found other cell types, and also in unicellular organisms, meaning that they were already present before the origin of multicellularity (Goodwin et al., 1993; Newman and Bhat, 2008; Newman, 2011). Thus, these easy-to-arise patterns may have represented the raw material upon which evolutionary forces may have acted in order to build more complicated patterns later on (by using more cell processes and regulating their spatial and temporal location finely).

Besides development, that by means of combined cell processes specifies which patterns are more likely to appear, selective forces should also contribute to understand why some of the cleavage patterns are evolutionary conserved while others not. For some authors, the maintenance of some cleavage types within major invertebrate clades suggest that there is an important reason for them to be conserved (Valentine, 1997). As in other evolving biological structures, the adaptive significance of these conserved patterns may rely in two (not exclusive) facts:

First, if cell fates are specified by cell-cell interactions between specific blastomeres, a constant relative position between these blastomeres is crucial for the appearance of functional adult organs (Salazar-Ciudad, 2010). In these cases, variations in the blastomere positions within the blastula should be maladaptive and selectively suppressed. This is often referred as “internal selection” (Riegler, 2008).

Second, adaptive modifications of some aspects of early development (e.g. an increase in the amount of yolk in order to nourish the embryo, or a hardening of the eggshell) may, in turn, have an effect on the blastomeres´ shape and arrangement (Wray, 2000).

3. Evolution of cleavage in metazoans.

In order to see how the described phenomena can account for the different cleavage patterns and their evolutionary transitions, we present an overview of the cleavage patterns among the extant metazoans. Unlike other reviews of early development of metazoans, the one we present here is exhaustive (it covers at least all the highest taxonomical levels of all known animal phyla), comprehensive (such a general approaches have been normally published as books) and, more importantly, based in the cell processes involved in each cleavage pattern. We have considered the main taxonomical groups as are widely recognized in recent literature (Ax, 2012; Nielsen, 2012), without adhering to any particular phylogenetic hypothesis.

3.1. Non-bilaterians.

Despite their apparent morphological simplicity, the cell processes and the cleavage patterns deployed during the early development of these groups (Poriferans, Placozoans, Cnidarians and Ctenophores) are extremely diverse (Adamska et al., 2011). In general, they show holoblastic cleavage patterns (even though some species have abundant yolk) that are characterized by their irregularity. The direction of cell divisions is random (so that each blastomere has its own axis of polarization and cleaves independently of the others) and cell divisions become asynchronous soon after fertilization, resulting in amorphous blastulae with low cohesion and no recognizable geometrical regularities. This cleavage type is called anarchical or chaotic. In many species, specially among cnidarians, the cleavage pattern is not only chaotic but also variable between individuals (involving even transient syncytial stages by random fusion between blastomeres or by anomalous cytokinesis). As a result, the number of chaotic cleavage patterns is very large (its sometimes impossible to find at least two egg cells with a similar position of the blastomeres, even in the same species (Kauffman, 2004)).

This great spatio-temporal variability of the chaotic cleavage pattern (even if it presents a transitory regular appearance) prevents the determination of the cell fates during early cleavage: the cell fate of each blastomere can not be unequivocally determined by its embryological context, since the relative position and identity of its surrounding blastomeres is far from constant (Freeman, 1983; Kauffman, 2004). This has been demonstrated in hydrozoans, in which normal development is not affected by the artificial displacement of blastomeres (Freeman, 1983; Kauffman, 2004). Notwithstanding this, not all non-bilaterians exhibits the same degree of chaoticness in their early development, and each phylum deserves a more detailed explanation:

3.1.1. Porifera

Sponges are one of the simplest animals, and probably the taxon that branched from the rest of the metazoan lineages earlier (Nielsen, 2012). They are composed of very few (seven) types of cells. Based on differences of the topology of their filtering system and their biochemical composition of their skeleton and spicules, four main classes (*Calcarea*, *Hexactinellida*, *Demospongiae* and *Homoscleromorpha*) are normally recognized (Brusca and Brusca, 1990; Gazave et al., 2012). Despite the early development of poriferans is of interest because of their phylogenetic position among metazoans, studies on this topic are scarce (Ereskovski, 2002; Adamska et al., 2011), and even more scarce for oviparous sponges (Leys and Ereskovski, 2006). In non-oviparous species cleavage proceeds inside a special brood chambers. In poriferans, eggs are rich in yolk inclusions evenly distributed in the cytoplasm (isolecithic), but the cleavage is still holoblastic.

A variety of cleavage patterns can be distinguished, being the chaotic the more predominant (it is found in all *Homoscleromorpha* and most ovoviviparous *Demospongiae*). During chaotic cleavage, cell divisions are often both asymmetric and asynchronous (like in the model system *Amphimedon queenslandica*), thus being the blastulae composed of irregularly shaped macromeres of various sizes with small micromeres interspersed between them (Adamska et al. 2011; Ereskovski, 2013). During these disordered cell divisions, some poriferan taxa show transitory ordered patterns resembling those found in spiralians and deuterostomes (see Section 3.4 and 3.5). These are called, respectively, pseudospiral and radial-like patterns.

In the former (that is found at least in the genus *Oopsacas* (*Hexactinellida)*) the pattern up to the 8 cell-stage is alike to quartet spiral pattern (see section 3.4) but in this case, the relative twist between micro- and macromeres in the 8-cell stage is variable within a species. Moreover, the pseudospiral pattern vanishes in further stages, when blastomeres start to divide asymmetrically, and the resulting micro and macromeres sort the way that yolk-rich macromeres lie at the center of the blastula surrounded by an external layer of micromeres (Leys and Ereskovski, 2006). These peripheral micromeres are connected by cytoplasmic bridges and eventually fuse to form a perypheral syncytium. The possible evolutionary relationship between these two patterns (transition from pseudospiral to spiral cleavage patterns) may rely on whether the true spiral pattern is mainly produced by a rearrangement of blastomeres (as is the case for the pseudospiral pattern) or it requires a finer regulation by the ooplasm (Kauffman, 2004).

In the radial-like pattern (showed by oviparous *Demospongiae* and some *Hexactinellida*) the two first cell divisions are meridional, and the third one is equatorial. This generates eight

blastomeres in a cubic configuration (each one placed in the corner of an hypothetic cube). However, subsequent cell divisions are, contrary to deuterostomes, all perpendicular to the surface of the embryo, generating a hollow blastula of elongated cells (Ereskovski, 2006). According to some authors (Watanabe, 1978), the direction of these early cell divisions are dictated by the point of adherence of the embryo to the substrate, at least in the Demospongiae family *Tetillidae*.
Thus, these ordered patterns are restricted to the very early stages (before 8-cell stage) and appear only in some individuals within a species, suggesting that they are likely to be produced by the mechanical stability of cell adhesion between blastomeres (best packing configurations).

Some sponges also exhibit truly non-chaotic and non-transitory patterns. In *Halisarcida*, the most primitive order of *Demospongiae* (Ereskovski, 2002), each cell division is perpendicular to the cell surface and its direction is independent respect to the neighboring cells (polyaxial cleavage, see Fig. 1A). This results in a hollow blastula of cone-shaped blastomeres (with the cone tip pointing towards the blastula center). Some authors argue that this mode of cleavage may represent an evolutionary intermediate stage between chaotic and radial cleavage patterns (Ereskovski, 2002), but if this is actually the case is not clear.

In the genera *Leucosolenia* and *Sycon* (class *Calcispongiae*, subclass *Calcaronea*) all cell divisions are perpendicular to the animal-vegetal axis (incurvational cleavage, also called “table palyntomy”) (Ereskovski, 2013). This process generates a epithelioid layer of cyclindric blastomeres that bends in a cup-shaped blastula (being the concavity pointing towards the animal pole). This bending may be explained by the limited space available inside the incubation chambers of these species, that prevents the unconstrained planar growth of the blastula. Within the plane formed by the blastomeres (that is, when the blastula is viewed from the animal pole), direction of cell divisions are not random, but they form a tetra-radially symmetric structure (Anakina, 1997). This suggest that, besides planar polarity, other mechanisms determining the direction of cell division, such as Hertwig's rule or Sachs' rule should exist.

In poriferans, cleavage normally finishes in a blastula-like spheroidal aggregate of ciliated cells, that is actually the swimming larva. Depending on whether the aggregate is hollow or not, on its overall shape and on the presence and location of ciliated cells, eight types of poriferan larvae are defined (Leys and Ereskovsky, 2006).

3.3.2. Cnidaria

The main representatives of cnidarians include hydrozoans (hydra-like animals), anthozoans (corals and sea anemones), scyphozoans (jellyfishes) and cubozoans (box jellyfishes, sea wasps). Their eggs have yolk and lipidic substances for nurture and, in sessile species such as anthozoans, also for floating. In this latter case, the ability to float (and disperse) depends not only on the amount of these fatty substances but also on its wax ester content. These lipid droplets are small at early stages but merge together during gastrulation, thus forming larger aggregates (Daly et al., 2013; Okubo, 2013).

Cleavage is holoblastic, but in the first cell division of species with yolk-rich eggs, the furrow has some difficulties in traversing the whole cytoplasm, producing a hearth-shaped or bean-shaped blastula (Okubo, 2013). The location of this first cleavage furrow correlates with the invagination site of the larvae, at least in some species (*Nematostella*, anthozoan) (Fritzenwanker et al., 2007).

Cnidarians have a very chaotic cleavage patterns (Fig. 1B), including asymmetric cell division as in the genera *Obelia* and *Aglantha* (Greenberg, 1959; Kauffman, 2004), random fusion between already formed blastomeres as in *Nematostella* (Fritzenwanker et al., 2007), transitory syncytial stages (when karyokynesis takes place without cytokynesis) as in *Obelia* and even the partial disintegration of the blastula as in *Turritopsis* and *Oceania* (Kauffman, 2004).

As in poriferans, some cnidarian species (e.g. *Rathkea*, hydrozoan pseudosp.) show transitory ordered patterns. Radial-like or pseudospiral configurations have been reported for some coral species and the hydrozoan *Nematostella* (Kauffman, 2004; Fritzenwanker et al., 2007; Okubo, 2013). However, as in poriferans, these non-chaotic patterns exhibits great intra-specific variation and they are not detected detected after the 8-cell stage (Scholz et al., 2003; Fritzenwanker et al., 2007; Okubo, 2013). Thus, they are likely produced by means of passive packing forces (cell-cell adhesion), although some kind of Sachs' rule like mechanism can not be discarded).

Interestingly, the same degree of chaoticness of cnidarian embryos correlates with their life cycles: the more important in postembryonic development the sedentary stages, the more anarchic the cleavage and the weaker of connections between blastomeres (Kauffman, 2004). That is, the cnidarian species that lack or have a weakly developed sedentary polyp stage in their life cycles (or they have lose it) exhibit more ordered cleavage patterns. This is the general case in Scyphozoans (e.g. *Pelagia*, *Chrysaora*) and some hydrozoans (*Aglantha*, *Aequarea*, *Amphisbetia*), in which the cleavage pattern follows a much less variable radial-like arrangement, at least in its first stages (Mergner, 1971; Kauffman, 2004). This is an example of how adaptation to different life-history

strategies may eventually drive the presence of one or another cleavage pattern (Wray, 2000).

Phylogenetic studies show that, within cnidaria, this radial-like patterns are evolutionarily original, being the more chaotic ones a derived type (Kauffman, 2004), but concluding evidence is still missing.

In cnidarians, cleavage finishes in a hollow sphere of cells (coeloblastula) that (contrary to bilaterians) gets invaginated in the animal pole, generating a cylindric, ciliated and swimming larvae called planula (Brusca and Brusca, 1990). Different cnidarian species undergo this gastrulation process in different modes (actually they display all modes of gastrulation known for metazoans: invagination, unipolar and multipolar immigration, epiboly and cellular delamination) (Okubo, 2013; Technau et al., 2015).

3.3.3. Placozoa

Placozoans are the simplest metazoans from a morphological perspective. They basically are composed of two cell layers (basal and dorsal epithelia) with some other sensitive cell types intercalated (Brusca and Brusca, 1990). Despite placozoans (*Trichoplax adhaerens* and few more unnamed and morphologically identical species (Eitel et al., 2013; Schierwater and Eitel, 2015)) have been studied since a century ago, their life cycle was only described in 2011 (Eitel et al., 2011), well after their genome was sequenced (Srivastava et al., 2008). Placozoans can undergo asexual (by budding) or sexual reproduction. In this latter case, oocytes are thought to derivate from the lower epithelium (Eitel et al. 2011), and have big droplets of “fatty substance” forming cytoplasmic accumulations and/or extracellular deposits. However, when placozoans are induced to follow sexual reproduction under laboratory conditions, embryos never develop beyond 64 cell-stage (the reason for that is still debated (Eitel et al., 2011)).

Before this 64-cell stage, the cleavage is total (holoblastic), synchronous, equal (all blastomeres are equally sized) and chaotic. In the 4-cell stage, blastomeres are placed in a tetrahedral configuration, which is known is the optimal (less energy) packing configuration of four spheres (Manoharan et al., 2003). This is compatible with the eggshell imposing some compressive effect (cleavage takes place inside a protective eggshell: the fertilization membrane).

After the 64 cell-stage, few in vivo observations (Grell, 1971) let us know that cleavage continues being synchronous, total and equal at least in the 128-cell stage and 256-cell stage (Schierwater and Eitel, 2015).

3.3.4 Ctenophora

Ctenophores, also known as comb jellies, are marine animals whose evolutionary affinities have been debated for many years (Brusca and Brusca, 1990; Rupert and Barnes, 1994; Nielsen, 2012). Their most distinct feature is the presence of large bands of cilia, that form a swimming structures called combs or comb plates. Along with a pair of tentacles (absent in some groups) and sensory organs, these structures are organized in a radial manner around the main body axis: the oral-aboral axis. That way the adult ctenophore body is composed of four nearly identical quadrants organized around this axis, producing a tetra-radial symmetry (Henry and Martindale, 1999). Each one of these quadrants is largely derived from one cell of the four-cell-stage embryo. Thus, since ctenophores develop rapidly and have not larval stages, its symmetry axes relate from the very beginning with those found in the adult organism (Freeman, 1977), which has lead to some authors to argue that factors specifying the location of some adult body structures (e.g. comb plates) are already localized in the ctenophore egg, which implies some kind of premorphological organization (Driesch and Morgan, 1895; Freeman, 1977).

The knowledge of the ctenophoran early development comes from the model species *Mnemiopis leidyi*. In this species, and contrary to other bilaterians, the cleavage is very regular and exhibits no variability. Concerning the spatial distribution of the blastomeres, this pattern does not resemble any other one found in metazoans, which do not help to clarify the phylogenetic position of this controversial taxon. The eggs have yolk settled in the vegetal hemisphere, but the cleavage is holoblastic, with the furrow of the first cell division originating in the site where the polar bodies were budded. This site corresponds to the future oral region of the larva (thus, the oral-aboral axis is established by the first cell division). Second cell division occurs perpendicular to the first one, but also the same oral-aboral axis, yielding four equally sized blastomeres arranged in the same plane (their cell fates, however, are not the same, see below). In the third cell division, that is oriented obliquely to the oral-aboral axis, four new cells appear closer to the oral pole, but displaced towards the periphery of the blastula, forming a “C”-shaped blastula in lateral view (The opening of the “C” corresponding to the aboral pole and the concavity to the oral one, see Fig. 1C). When viewed from the aboral pole, this 8-cell stage blastula is not radially but bilaterally symmetrical: the four aboral-most blastomeres (called E-blastomeres, that are slightly smaller than those located in the oral pole, called M-blastomeres) contact between them two by two. In this view, the extremes of the blastula where the sets of two contacting E-blastomeres are located define the tentacular plane (where the tentacles will bud off). The perpendicular plane to this one, where no contacting E-blastomeres are found, define the sagittal (aka esophageal) plane. Thus, the tetraradial symmetry of this organism is

defined in the very early cleavage. This cleavage pattern is called biradial by some authors, but it should not be confused with the also called biradial pattern present in Lophophorates (see section 3.4.2.1) (Actually, according to its symmetry axes it would be also called bilateral (as in ascidians, see Section 3,4,6), but to our knowledge this term has not been used in literature.

In the next cell division, each blastomere gives rise to two cells, a macromere and a much smaller micromere. This division is oriented, for all cells, towards the aboral pole of the blastula. As a results, all the 8 new micromeres are located in the blastula concavity, forming a ring. Microinjection experiments have revealed that the micromeres that are located the furthest away from the blastula center will give rise to the comb plates (Freeman, 1977; Martindale and Henry, 1999). In the next two cell division rounds (32 and 60-cell stages), aboral micromeres continue dividing along the border of this aboral ring-like structure, whereas the macromeres continue dividing asymmetrically towards the aboral pole. At the 60 cell stage (the four bigger and oral-most macromeres do not divide in the 6^{th} division round), the blastula is composed of eight large, yolk rich cells in the oral pole, that will generate both endodermal and mesodermal cell types, and 52 much smaller cells between them that will give rise to the ectoderm. Soon after this stage, gastrulation takes place by epiboly: aboral micromeres spread over the larger macromeres toward the animal (oral) pole. During gastrulation, a secondary set of tiny micromeres (also mesodermal precursors) are given off by the macromeres at the oral pole (Ortolani, 1963; Freeman, 1977).

Concerning the mode of cell fate specification, ctenophorans (as other metazoans with a constant cleavage programs) have precocious specification of cell fates by the segregation of morphogenetic factors. These factors, that are unevenly distributed within the blastula, are segregated in specific blastomeres by coordinated asymmetric cell divisions. However, this non-uniform distribution of the factors does not exist before fertilization; if an unfertilized ctenophore egg is cut in two halves and then fertilized, each one develops into a normal adult. Some factors are heterogeneously distributed as a result of the first cell division: if the site of this cell division is experimentally changed (by centrifugation), the embryonic structures appear in the new relative orientation, not the old one. Other factors are actively segregated to the aboral micromeres at the eight-cell stage (Driesch and Morgan, 1895; Spek, 1926). Thus, the cleavage itself is causally involved with highly coordinated segregation of morphogenetic factors (Freeman, 1977), which in turn are used for cell fate determination (although cell-cell induction is also utilized in later developmental stages) (Martindale and Henry, 2015).

3.2. Xenacoeloelomorpha

Xenacoelomorpha (Acoela, Nemertodermatida and Xenoturbellida) are thought to branch from the rest of the bilateria very early on. In general, their cleavage is holoblastic, and shares some characteristics with the spiral one (Wanninger, 2015). Despite the recent advances in the biology of the enigmatic phylum Xenoturbellida (Rouse et al., 2016), embryological studies have not been conducted so far (Hejnol, 2015).

3.2.1. Acoela.

Because of their simple organization, these small flatworms have played a prominent role in discussions regarding the origins of bilaterian phyla. They have been classically considered to be descendants of coelomates (platyhelminthes) through secondary loss of derived features (Brusca and Brusca, 1990), but more recent phylogenetic studies using 18S rDNA suggests that they may be the earliest divergent bilaterian taxa (Ruiz-Trillo et al. 1999). This controversy about the phylogenetic position of acoel flatworms is still held (Egger et al., 2009). In this evolutionary debate, the cleavage type exhibited by acoels (a special, taxon-specific pattern called duet spiral) is considered to be relevant (Jenner, 2004).

In the duet spiral cleavage pattern, first cell division proceeds meridionally, and the plane of this first cleavage corresponds to the plane of bilateral symmetry in the adult body (Henry et al., 2000). Second cell division, in contrast to the quartet spiral cleavage, is not meridional but oblique respect to the animal-vegetal axis, and directed towards the animal pole of the blastula. This produces two micromeres placed over the two macromeres (Fig. 1D). These two micromeres contact each other as the two macromeres do, but in a plane that is rotated around the animal-vegetal axis (Henry et al., 2000). This means that when viewed from the animal pole, the pair of two micromeres seem to have rotated (counterclockwise) respect to the underlying pair of macromeres. The next generation of micromeres also appear in the central part of the blastula close to the animal pole, but they have an apparent contrary (clockwise) rotation when viewed from it. When macromeres divide in the third cell division, they do in the direction of the animal-vegetal axis. In general (see Henry et al., 2000 for the genus *Neochildia* and Brusca and Brusca, 1990 for the genus *Polychoerus*) all blastomeres in the 4-cell stage does not divide synchronously, but micromeres divide first and macromeres later, thus passing through a transitory 6 cell-stage. In the fourth cell division round (8 to 16-cell stage), the most vegetal macromeres (2Q) divide first (symmetrically and meridionally), giving rise to a row of four macromeres in the vegetal pole of the

blastula (the two central ones, 3A and 3B are endomesodermal precursors that will become internalized during gastrulation). The rest of the blastomeres divide also in a symmetric manner but following different spatial directions: approximately, micromeres (1q1 and 1q2) divide along the animal-vegetal axis (equatorially), whereas the (2q) macromeres do it meridionally. Soon after this 16-cell stage, gastrulation begins with the internalization of the endomesodermal precursors 3A and 3B (Hejnol, 2015).

In some of these early stages, and in all species investigated so far the blastula is reminiscent of the quartet spiral one, but laterally compressed. Actually, even the Conklin nomenclature, a standard labeling for homologous blastomeres in the quartet spiral pattern (see Section 3.4, Gilbert and Raunio, 1997) has been used (with minimal adaptations consisting in the “deletion” of the C and D cell lineages) for the labeling of blastomeres in this this duet spiral cleavage (Hejnol, 2015).

Mechanistically, it seems that the main difference to the quartet spiral cleavage (see section 3.4) relates to the timing in which the cell processes are deployed: in acoela, the “spiralizing” events leading to oblique cell divisions start one cell-cycle earlier (in the 2-cell stage) and the synchrony between cell divisions is lost earlier than in Spiralia (Henry et al., 2000). The mechanisms specifying the clockwise-counterclockwise alternation are also different, but it is conceivable that, having the same mechanism (e.g. Sachs's rule), the outcomes are different because of the different spatial arrangement of blastomeres, or because acoel embryos skip some “even” cell divisions (Henry et al., 2000).

The cell fates and modes of cell fate determination also differ substantially between spiralia and acoela. In quartet spiral pattern cell fate determination is highly determinative (autonomous determiantion of cell fates), but cell deletion experiments show that in the case of the duet spiral cleavage development is highly regulative (inductive determination of cell fates) (Boyer, 1971). The fates of individual blastomeres are also not the same. Despite cell tracking experiments beyond the 12-cell stage cell stage are technically difficult (Henry et al., 2000), both the orientation of early cleavage planes and their relationship with adult axes differs substantially between duet and quartet spiral cleavers (Henry et al., 2000).

Thus, these extensive differences suggest that the duet spiral cleavage pattern is not a derived form of the quartet spiral cleavage, but that both quartet and duet spiral cleavage patterns have likely evolved independently from an ancestral radial-like cleavage pattern (Ruiz-Trillo et al. 1999). However, a deeper understanding of the biomechanics of the duet spiral cleavage would be required in order to clarify whether it can be easily derived from a radial-like pattern or not (because so far, the presumed phylogenetic position of acoela has had a deep effect upon the

interpretation of their cleavage pattern (Jenner, 2004)). This, in turn, may help to clarify the evolutionary origins of this invertebrate taxon.

3.2.2. Nemertodermatida.

As acoel flatworms, the nine species of this phylum are small marine worms that are considered to belong to one of the earliest bilaterian lineages (Jondelius et al., 2002). The phylogenetic relationships between these two lineages are still discussed, and some authors classify them within a single clade called Acoelomorpha (Ruiz-Trillo et al., 2002). Since cleavage patterns of embryos often provide important cues for higher level metazoan phylogenetic relationships (Jondelius et al. 2004) similarities in the early development of acoela and nemertodermatida would favor the inclusion of both within Acoelomorpha, but recent reports of early development in nemertodermatida dismiss this hypothesis (Jondelius et al., 2004; Børve and Hejnol, 2014; Hejnol, 2015).

In the model species *Nemertoderma westbladi* eggs are quite rounded and they are covered by an eggshell. Cleavage is holoblastic, being the first cell division symmetric and meridional (the animal pole marked by the polar bodies). Second cell division is asymmetric and equatorial, lying two smaller cells (micromeres) close to the animal pole (as in acoel worms). After cell division, the two micromeres shift slightly clockwise when seen from the animal pole (contrary to acoels, which do it counterclockwise). Next cell divisions is asynchronous (micromeres does not divide as the same time as macromeres) and (also contrary to acoels) meridional, so that the 8-cell stage closely resembles the 8-cell stage in the quartet spiral cleavage.

The next cell division is also meridional, thus being different from quartet or duet spiral patterns. In this cell division, that is meridional for all cells, the cell divisions planes pass through the center of the blastula. As a consequence, in the 16-cell stage the eight macromeres form a rosette around a vegetal pole and the eight micromeres do the same in the animal pole. Thus, from the animal pole the blastula presents an octo-radial symmetry. This special cleavage type has been described as unequal latitudinal cleavage (Jondelius, 2004).

Other species (*Meara stichopi*) also starts cleavage following a duet spiral pattern but micromeres in the 4-cell stage shift only very slightly counterclockwise (Børve and Hejnol, 2014). As a consequence, in the 8-cell stage the four animal blastomeres are situated directly on top of the vegetal blastomeres and not between the vegetal blastomeres as it is the case in spiralian embryos. The next cell divisions are symmetric and rather asynchronous, but the arrangement of blastomeres in further stages does not seem to follow a regular pattern. At 24-cell stage a blastocoel appears, and

gastrulation takes place at about 64-cell stage by the internalization of some blastomeres (presumably endomesoderm precursors) into the blastocoel. The detailed cell fate map for Nemertodermatida have not been established so far.

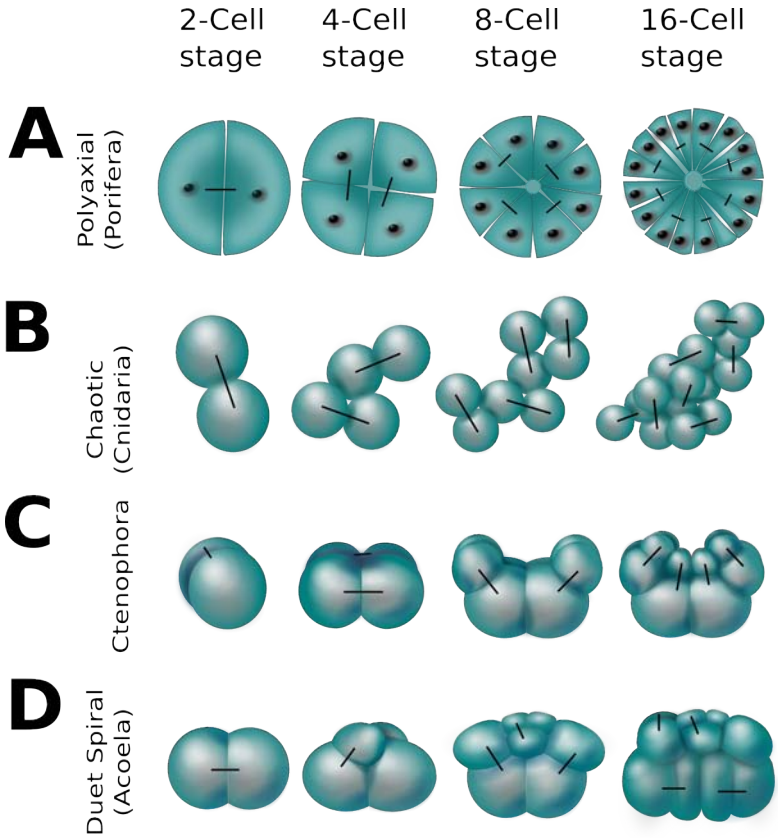


Figure 1. Example cleavage patterns found in non-bilaterian metazoans and acoel flatworms. Most of them may be explained by means of the combination of a few conserved processes (see text). All blastulae are displayed in lateral view with the animal pole on the top, and the small straight lines link sister blastomeres when both of them are visible and identifiable. A) Polyaxial cleavage of some poriferans (*Demospongiae*). Drawings are idealized sections of the blastulae. B) Chaotic (=anarchic) cleavage pattern characteristic of non-bilaterian taxa such as Cnidarians.) C) The cleavage pattern of Ctenophores. D) The duet spiral pattern of acoel flatworms.

3.3. Ecdysozoa

Ecdysozoa is a vast and recently established superphylum of animals characterized by the presence of an external cuticle and their ability to moult (Aguinaldo et al., 1997; Telford et al., 2008; Telford et al., 2015). Their huge diversity in the developmental modes of ecdysozoa (and in all other aspects of their biology) deserves a detailed description a of each group:

3.3.1. Scalidophora.

These three small phyla (*kynorhyncha, priapulida, loricifera*) of marine invertebrates form the clade Scalidophora, which supported by the fact that both of them have chitinous cuticle that is moulted, a ring of spines (scalids) in the introvert, and two rings of introvert retracts (Brusca and Brusca, 1990; Ax, Dunn et al., 2014). All of them are direct developers: their larvae have the same morphology as adults, in miniaturized scale (Hejnol, 2015).

Nothing is yet known about the early development of loriciferans (Brusca and Brusca, 1990; Hejnol, 2015), and for the kynorhynchs, mating has never been observed, and both egg laying and early development (including cleavage patterns) have not been adequately studied (Brusca and Brusca, 1990). Limited evidence from *Echinoderes kozloff* shows that cleavage is equal and almost totally synchronous, and that blastomeres are loosely joined (Kozloff, 2007). First cell division is meridional and the direction of the second one is not well defined. In the four-cell stage blastomeres have a tetrahedral arrangement (most likely because this is a very stable configuration (Song et al., 2008)), and by the 8-cell stage the embryo lacks a recognizable pattern. After the 16-cell stage, some blastomeres are internalized by means of an unknown mechanisms (candidates include the packing constrains imposed by the eggshell: simply not all dividing blastomeres can occupy peripheral positions). Despite this fact, the pattern has been described as a radial-like one or even as similar to priapulids (Hejnol, 2015). Because of this internalization, blastula gets organized into concentric blastomere “layers”, without any landmarks that would indicate an orientation of the differentiation of cells. During all the cleavage process, the embryo is compact and does not seem to possess a blastocoel.

In Priapulids, the only scalidophoran taxon whose early development has been accurately described (at least in two of its twenty species: *Priapulus caudatus* and *Halycriptus spinulosus*), the cleavage pattern is holoblastic, synchronous, and sub-equal (slightly different sizes between micro- and macromeres) (Hejnol, 2015). The whole cleavage process takes place inside two membranes: a robust eggshell devoted to mechanical protection and a fertilization membrane (Wennberg et al.

2008). Up to gastrulation, cell divisions tend to occur at right angles to each other (stereotypical cleavage), generating a symmetric pattern that resembles the radial one (Fig. 2A). The two first cell divisions proceed meridionally, and produces four equally sized blastomeres (Wennberg et al. 2008; Hejnol, 2015). A cross furrow (contact zone between opposing blastomeres) is observed in both animal and vegetal poles. In *P. caudatus*, these cross-furrows are at right angle between them (if this is also the case in *H. spinulosus* is currently unclear (Wennberg et al. 2008). The next cell division is equatorial and slightly asymmetric, and thus the 8-cell stage contains four animal and four slightly larger vegetal blastomeres (Wennberg et al. 2008). All cell divisions from this eight-cell stage are symmetric and equatorial, giving rise to a 16-cell stage blastula that exhibits four rows of four blastomeres each aligned along the animal-vegetal axis (in each of these rows, the blastomeres closer to the vegetal pole are slightly larger).

After the 16-cell stage, the blastula is so compact (either by increased cell adhesion or compression from the eggshell) that the visible face of each blastomere acquires a polygonal shape, and the directions of cell divisions differ between the blastomeres (Hejnol, 2015). The vegetal-most and the animal-most blastomeres (which show triangular visible faces) divide along the meridional axis, while the remaining blastomeres (which show rectangular visible faces) divide in equatorial direction. Thus, the resulting 32-cell stage is again composed of four blastomeres with triangular visible faces cells in the animal and vegetal poles, and several rings of rectangular-shaped blastomeres between them. Again, the directions of further cell divisions seem to depend on these shapes: cell divisions take place along the longest axis of the visible face of each blastomere (this is specially clear for “rectangular” blastomeres, which again divide meridionally). This may imply that a Hertwig-like rule restricted to the outer blastomeres' faces is the main driver of priapulid cleavage.

After this 64-cell stage, gastrulation starts with the internalization of the vegetal blastomeres (endodermal and mesodermal precursors) by cell migration (Wennberg et al. 2008; Martín-Durán et al. 2012). In later gastrulation, when a secondary opening (the prospective mouth, as in deuterostomes) appears, the blastula looses its radial symmetry and becomes bilateral (Martín-Durán and Hejnol 2015).

According to Hejnol (2015), the broad diversity of cleavage patterns and developmental modes of the scalidophorans and the two following phyla (nematoda and Nematomorpha, all of them forming a super-clade called Cycloneuralia) make it difficult to reconstruct the developmental aspects of their last common ancestor. However, it seems likely that a radial cleavage pattern with deuterostomic gastrulation was part of the ground pattern of cycloneuralia.

3.3.2. Nematoda.

Nematodes or round worms comprise more than 25000 (Hodda, 2011; Zhang, 2013) morphologically similar species that inhabit a extremely broad range of environments. More than a half are parasitic, and they exhibit a diversity of life cycles. Nematodes usually reproduce sexually, and the eggs are protected by an chitinous outer shell (or eggshell) secreted by the uterus. The best studied nematode is the model organism *Caenorhabditis elegans*, which has played a pivotal role in the elucidation of basic principles of developmental biology. In this species, (and in most nematode species) the number of somatic cells in the adult body is known to be constant (a phenomenon known as eutely), and the cell lineage, that is almost entirely invariant between individuals, is specified from the very early developmental stages (Gilbert and Raunio, 1997; Somer, 2015). The cell lineages and arrangement of blastomeres in other nematode genera such as *Ascaris* are similar to the pattern found in C. elegans, which has lead to the assumption that embryonic development shows little variation within the phylum Nematoda (Voronov, 2000).

However, analysis of other, less known nematode species show that important variations in early embryogenesis exist (especially within the class Enoplia)), which is in contrast with the relative uniformity of the adult morphology (Goldstein, 2001; Schierenberg, 2001; Schierenberg, 2006). This variable cleavage is thought to be plesiomorphic within nematoda (Somer, 2015). In this review, we focus on the predominant invariant mode (specifically the well characterized cleavage of *C.elegans*, which is said to be holoblastic “rotational cleavage”) but highlighting, if proceeds, the main differences with other nematodan taxa.

As in many other nematodes, the early development of *C. elegans* takes place inside the maternal body. Eggs are fertilized with sperms stored in the spermatheca and they pass through a “U”-shaped uterus as they develop, until they are by spawned by the vulva (Gilbert and Raunio, 1997). The eggs are ellipsoidal, and its longest axis corresponds to the future antero-posterior (AP) axis. Which part of this axis will become the anterior part of the body is determined by several mechanisms depending on the species.

In some species (e.g. *C. elegans*), it depends on the point where the sperm enters the oocyte cytoplasm. The sperm pronucleus interacts with a family of subcortical proteins (PAR proteins) so that the cell cortex is no longer symmetric along the antero-posterior axis (Goldstein and Hird, 1996). This produces a differential attachment of spindle microtubules along this axis (stronger in the anterior half of the embryo), that results in a displacement of the mitotic apparatus towards the anterior part, and asymmetric cell division (Gilbert and Raunio, 1997; Goldstein, 2001). The cell asymmetries generated by the sperm entrance also involves the rearrangement of the so called P

granules (ribonucleoproteins involved in the germ line specification) in the posterior end of the embryo by means of active cytoplasmic movements.

In other species, especially those which develop from a partenogenetically activated eggs, the antero-posterior embryonic polarity can not be established by the sperm pronucleus. In these species, such as *Acrobeloides nanus*, the polarity of the embryo depends on the relative position of the egg within the uterus, and (with few exceptions) the anterior part points toward the vulva. However, the exact nature of these cues not involving the sperm pronucleus is not clear (Goldstein, 2001). Finally, in other species (e.g. *Diploscapter coronatus*) the anterior pole does not correlate with any spatial cue, so it seems to be established at random (Lahl et al., 2006).

Regardless the mechanisms involved in the AP polarity, the first cell division occurs along the AP axis. In most cases, the resulting anterior blastomere (AB, also termed founder cell) is larger than the posterior one (P1, a stem cell that contains the P-granules, and is the germ-line precursor) (Strome et al., 1994; Gilbert and Raunio, 1997; Seydoux and Schedl, 2001). There exist inter-specific variability concerning the degree of asymmetry of this first cell division: in *Trichuris muris* (clade I) the first division is extremely asymmetric (Malakhov, 1994) while in *Prionchulus* (clade I) asymmetric cell divisions can only be detected at a later stages. In clade II cell divisions are always symmetric, and no AP polarity has been detected (in these cases the germ-line precursors are thought to be generated via a different mechanism (Malakhov, 1994; Voronov, 1999; Lahl et al., 2006; Schierenberg, 2006).

In their predominant mode of cleavage (holoblastic rotational), the following cell division are generally symmetric. In *C. elegans*, during the second cell division, the founder cell (AB) divides equatorially (90 degrees to the antero-posterior axis), while the P1 divides asymmetrically and meridionally to produce another founder cell (EMS) and a smaller posterior founder cell (P2), which now contains the P-granules (Gilbert and Raunio, 1997) (see Fig. 2B). In this 4 cell stage, the blastomere arrangement within the eggshell is normally diamond-shaped, but other, less abundant configurations also appear (Kajita et al. 2005; Voronov, 2000). The presence of one or another of these configurations seems to rely on small random differences of cell-cell adhesion and cell shape, and on the compressive effect exerted by the eggshell (Schierenberg and Junkersdorf, 1992; Kajita et al., 2005; Goldstein, 2001). This compressive effect is supported by the way the eggshell shape correlates with different blastomere configuration in different nematode taxa. Eggshells whose longest and shortest axes are relatively similar (such as in *C.elegans*) allow for a diamond-shaped configuration (Kajita et al., 2005) or for a row of four cells (Dolinski et al. 2001), whereas very elongated eggshells can only contain the four blastomeres in a linear row (Goldstein, 2001; Dolinski et al. 2001).

Several mechanisms determine the direction of the ensuing cell divisions until gastrulation. Hyman and White (1987) showed that in all blastomeres of the *C.elegans* cleaving embryo, just after cell division the replicated centrosomes separate and migrate to opposite sides of the nucleus, which causes the mitotic spindle to be set up at a right angle to the mitotic spindle of the preceding division (Sachs' rule). In the AB cell lineage, whose cells are located in the anterior half of the embryo, this results in equal and synchronous cell divisions along successively orthogonal axes.
Instead, cells belonging to the stem (P) lineage divides always asymmetrically, more slowly than the rest of the divisions of the embryo, and roughly along the same AP axis as the preceding division (Goldstein, 1995; Gilbert and Raunio, 1997; Kipreos, 2005). In each cell division, this produces an anterior founder cell and a posterior cell that will continue the stem cell lineage (that is, the posterior P2 cell in the 4 cell stage gives rise to the C (anterior) and P3 (posterior) blastomeres, and in turn P3 gives rise to D (anterior) and P4 (posterior) blastomeres) (Goldstein, 1995; Gilbert and Raunio, 1997). Although these cell divisions oriented along the AP axis are compatible with a polarizing gradient, experiments show that this is not the case. Instead, in these cells the centrosome and nucleus rotate as a unit through 90° just after cell division and centrosome migration. This rotation counteract the “orthogonalizing” effect of the centrosome migration, leaving the mitotic apparatus oriented in the same axis as the preceding cell division (Nigon et al., 1960; Albertson, 1984; Hyman and White, 1997). After some cell divisions, this would result in a line of cells, but the constraints of the eggshell skew them somewhat off this axis (Goldtstein, 1995). Experiments also show that, besides these mechanisms, the direction of cell divisions of some early blastomeres (EMS and E) are controlled by specific cell-cell contacts (EMS-P2 or E-P3 contact) (Goldstein, 1995). Finally, cortical rotation has also been observed during these very early stages, with potential causal effects on the sorting of cytoplasmatic elements responsible of the establishment of antero-posterior and dorsoventral polarity (Hird and White, 1993; Singh, 2014).

Thus, cell adhesion, Sachs´ rule, cortical rotation, the eggshell and specific cell-cell contacts controlling the direction of cell division are involved in the emergence of the nematode cleavage patterns. The variations in these patterns within nematoda are likely explained by variations in the strength and timing of these mechanisms.

Just after the generation of the P4 blastomere (the germ line precursor), as early as 24 cell stage, the gastrulation begins. In *C.elegans*, gastrulation takes place by an internalization of some blastomeres to the center of the embryo, without involving a true blastocoel (some authors consider the small space between migrating blastomeres as a transient blastocoel). First, E blastomeres (intestine precursor) migrates from the ventral side to the center of the embryo, then the P4 cell, located at the posterior end, do the same, being now located in the posterior-medium part of the

embryo. Finally, the cells of the M lineage (muscle, pharynx and gonad precursors) get interiorized from the anterior side and the C and D cells (muscle precursor) from the posterior side. After gastrulation, each blastomere roughly occupies the place where the body structures it will give rise to will be formed during organogenesis.

The described processes of gastrulation, however, differ in some enoplid genera (eg. *Tobrilus*) (Voronov, 1999; Schierenberg, 2006), in which as cleavage proceeds blastomeres are placed in the periphery of the blastula, giving rise to a true coeloblastula which gastrulates at 64 cell-stage.

As we can see, within nematoda there exists considerable variation in the spatial arrangement of blastomeres, as well as in the timing and the order of early cleavages (Skiba and Schierenberg, 1992). These variations in early development cause that the contacts between blastomeres (both temporally and spacially) differ between species. Since cell-cell interaction is a relevant mechanism for cell fate specification (besides autonomous mechanisms) there is also variation in the cell lineages, and in the mechanisms whereby they are established (Malakhov, 1994) (comparative studies have established three broad categories of cell lineages in nematodes). Strikingly, in some genera (*Rhabdias*) that alternate between free-living and infective generations (involving some morphological differences between them), the early embryonic cell lineages of each generation are also different (Spieler and Scherenberg, 1995).

3.3.3. Nematomorpha

Nematomorphs are a small phylum of parasitoid worms, phylogenetically related and morphologically similar to the more known and specious nematodes. Both are often placed together in a superphylum called Nematoida (Telford et al., 2008). In the few species of nematomorphs for which cleavage pattern has been described, it is holoblastic (total), equal (symmetric cell divisions) and highly variable (Malakhov and Spiridonov 1984; Hejnol, 2015; Wanniger, 2015). Because of this variability, several interpretations of the nematomorph cleavage pattern have been proposed (Jenner, 2004), including modified spiral cleavage (Brusca and Brusca, 1990), radial (Zrzavy et al., 1998), or modified radial cleavage (Valentine, 1997).

In *Gordius*, a model genus for nematomorpha, the first cell division is meridional, and the two nuclei (that are closer to the animal pole than to the vegetal pole) are slightly displaced from the midline when viewed from the animal pole (that is, the meridional plane that contains the two nuclei is not perpendicular to the contact surface between the two blastomeres). If this displacement is produced by cortical rotation or other mechanisms is not clear. In the four-cell stage, the

blastomeres can be in a planar diamond-shape or in a tetrahedral arrangement, and cell division rounds can be delayed, so that three-cell and five-cell stages are common (Meyer 1913; Inoue 1958, Hejnol, 2015). In some species, this 4-cell stage is reminiscent of those of various nematodes and gastrotrichs, in which two pairs of two cells are positioned in planes perpendicular to each other (Schmidt-Rhaesa, 1999; Jenner, 2004). In the 8-cell stage the blastulae often presents a transitory pseudospiral appearance that vanish in further stages. Because of this high variation in early development, the cell fate of each blastomere is not specified until later stages (Wanninger, 2015) making impossible the tracing of individual blastomeres (Hejnol, 2015).

3.3.4. Arthropoda.

Arthropods, with over a million species described, are the most successful metazoans in terms of adaptation to most of environments and the most diverse in terms of the number of species. All of them have a rigid exoskeleton made of chitin (which is periodically replaced by moulting), segmented body and jointed appendages (Grassé, 1949; Brusca and Brusca, 1990). They comprise four subphyla (plus some other extinct ones), namely hexapoda, chelicerata, crustacea and myriapoda.

The early development of arthropods (including the two related phyla Onychophora and Tardigrada) is very diverse, making difficult to extract commonalities, but always results in a similar antero-posteriorly segmented body pattern.

In general, the geometry of blastomere arrangement is fairly irregular and extensive cell movements are observed during cleavage, being groups of cells often shifted or substituted by adjacent cells. Thus, homologous structures can not be traced back to similarity of positions of the cells in the embryo, and fate maps can not be used to demonstrate phyletic relationships within arthropoda or between arthropoda and other metazoan clades (Weygold, 1979; de Beer, 1971; Scholtz and Wolff, 2013). In that sense, some purported cell-fate mapping in Arthropoda (Anderson, 1971) were proven to be a mere projection of later stages backwards (Scholtz and Wolff, 2013).
Because of that, most developmental studies carried out in these taxa are more focused on the shifting between high and low yolk content (its evolutionary drivers and developmental consequences) and on later developmental stages (germ-band stage, see below).

Arthropod eggs are normally elongated, although spherical eggs can be found in in some species of myriapoda and in chelycerates (Fig. 2C-D). Besides their asymmetry in shape, arthropod eggs are also asymmetric in cytoplasmatic composition (determinants or protein gradients that regulate the formation of anterior and posterior structures in further development. Examples of

these proteins are Bicoid (high concentration in the anterior part of the Drosophila embryo) and Nanos (high concentration in the posterior part of the Drosophila embryo), and the Nanos orthologs in non-dipteran arthropods (no Bicoid orthologs have been found outside diptera) (Gilbert and Raunio, 1997; Patel, 2000; Lall et al., 2003; Damen, 2007). These asymmetries, that are essential for further development, are usually maternally provided so that both antero-posterior and dorso-ventral axis are defined even before fertilization (Nusslein-Volhardt and Roth, 1989). An extra source of maternally inherited spatial information is the asymmetrically (to one side) positioning of the egg nucleus before fertilization. This is known to be important in the dorsoventral patterning of the embryo in insects (Lynch et al., 2010).

These eggs are very yolky and, in general, display meroblastic (incomplete) cleavage. Some authors speculate that this increase in the amount of yolk may have deleted all traces of spiral cleavage in the early evolution of arthropods, but this point is far from clear (Remane, 1983; Weygold, 1979). In the special meroblastic cleavage of arthropods, called intralecithal or superficial (Fig. 2D), non-cellularized nuclei start dividing deep within the yolk and then get displaced to the periphery forming a monolayer around the egg, that is known as periplasm (before cellularization) or more often, blastoderm (after cellularization) (Gilbert and Raunio, 1997; Chipman, 2015). This mode of cleavage, best exemplified in insects, is characterized by the absence of cytokinesis, which leads to the lack of cell membranes between the cleavage products, the so called energids (Scholtz, and Wolff, 2013). In other cases, the cytokineses are almost complete but the yolk and the blastomeres divide in one (2D) side of the embryo from the very beginning, so that the energids do not migrate from the center to the periphery (discoidal or superficial cleavage) (Scholtz and Wolff, 2013). In this case, and due to Sachs' rule, apparently regular (squares of 2 or 4 cells in each edge) configurations are often visible (Fig. 2C). Once the energids are in the embryo's surface (by means of cell migration or not), they get cellularized and the yolk remains in central position (centrolecithal cleavage). Notice that both nomenclatures (introlecithal and centrolecithal) simply describe the topological relationship between the yolk and the nuclei in two moments of the developmental dynamics of the arthropods.

The phase of cellularization is accompanied by a slowdown of nuclear divisions driven by changes in the chromatin/cytoplasm ratio, a transient stage that is called mid-blastula transition (Gilbert and Raunio, 2015). The relative time time of cellularization differs between different arthropod species (Nakamura et al., 2010).

There are a number of cases (see below) in which arthropod cleavage is holoblastic (Grassé, 1949). In these groups, the holoblastic cleavage is thought to be a side effect of an evolutionary loss of yolk that, in turn, is due to the non-lecitotrophic nutrition of embryos (e.g. viviparism or

planktotrophic larvae (Anderson, 1973)). Furthermore, in some special cases, total (holoblastic) and incomplete (intralecithal) cell divisions are combined different regions of the embryo or in different temporal stages (holoblastic cleavers which switch later in ontogeny to cleave superficially (to meroblastic cleavage), or vice versa) (Scholtz and Wolf, 2013). These modes of cleavage in which the holoblastic/meroblastic distinction does not hold for all cleavage process are referred as mixed cleavage (Dawydoff, 1928).

Concerning cell fate specification, determinate (cell types specified mainly by the cell lineage) and indeterminate (cell types specified mainly by positional clues in the blastoderm embryo) cleavages and forms that show aspects of both are found among arthropods (Gilbert and Raunio, 1997; Scholtz and Wolf, 2013). For those taxa that have determinate cleavages, the potential homologies between the cell fate maps of the different lineages are still debated (Peterson and Eernisse, 2001; Jenner, 2004).

Despite this review is mainly devoted to early cleavage, it worths mention a very important event in the gastrulation process of arthropoda: the germ band formation. The germ band is an elongated field of blastoderm cells lying at the ventral midline of the embryo that will give rise to the trunk of the embryo (the rest of the egg is normally formed by extra-embryonic tissue: yolk and the dorsal epithelium or amnioserosa) (Gilbert and Raunio, 1997). This implies that the ventral sides of the future metameric (repetitive) structures such as segmental furrows, have to be drafted in the germband. The germ band is generated by the migration and convergent extension of the ectodermal cells of the embryo surface, involving in general (insects, crustaceans, and arachnids) extensive cell movements. Comparative studies in insects have shown that there are two strategies whereby the germ band can generate serial structures (Krause, 1939), although intermediate strategies have also been described. These two strategies are called short and long germ-band. In short germ-band arthropods, the germ band initially comprises only the head lobes, and the rest of segments are successively budded by a posterior growth zone (generally in antero- posterior sequence) and differentiated one by one. In long germ-band arthropods, segments are formed almost simultaneously along the whole germ band (Gilbert and Raunio, 1997; Scholtz and Wolff, 2013). Some taxa display an “intermediate” germ-band formation in which the first segments are formed as in the long-germ band taxa, but the latter ones are added in a sequential manner, as in short-germ band taxa. The molecular mechanisms involved in each type of germ-band formation are different, but it has been shown that evolutionary shifts between them (long to short germ-band) are possible (Salazar-Ciudad at al., 2001; Damen, 2007)

The developmental stage in which the germ-band gets formed and segmented is shared by all arthropods and seems to be buffered against evolutionary change (Kalinka et al., 2010). This

taxon-specific conserved developmental stage is commonly referred as phylotypic stage (Sander, 1983).

Besides these general features, let us take an overview of the particularities of early development in each one of the four subphyla of arthropoda:

1) Hexapoda, including Insects and collembola (springtails):

In insects, eggs are produced by means of two different processes. In the first one, all the oocyte precursor cells differentiate directly into oocytes. In the second one, some oogonia differentiate into a specialized “nurse cells” which provide to the egg metabolic support and maternal determinants through active cytoskeletal transport (Cooley and Theurkauf, 1994). Most holometabolous (maximal indirect developers) insect have nurse cells, but the phylogenetic pattern seems to imply gains and loses of nurse cell during insect evolution (Chipman, 2015).

Its cleavage is in general meroblastic intralecithal with some variations in the time of cellularization. In holometabolous insects (e.g. *Drosophila*) cell divisions during cleavage are fast and the cellularization event occurs late, (well after the energids have reached the surface and have divided many times, generating about 5000 cells), whereas in hemimetabolous insects (e.g. *Grillus*, *Oncopeltus*, *Schistocerca*) cell divisions are slow and cellularization takes place before the blastoderm formation (Lynch et al., 2010; Nakamura et al., 2010; Chipman, 2015).
Holoblastic cleavage (described as chaotic by some authors, Anderson, 1973) is also found in a few pterygotes (some aphids and parasitic wasps), what is likely related to their derived reproduction modes (polymebryony, parthenogenesis, viviparism) (Weygoldt, 1979; Grbic, 2003; Machida, 2006). In hexapoda, all types of germ-band formation (short, long or intermediate) can be found (Sander, 1976).

2) Chelicerata, including arachnids, pycnogonids (seaspiders) and merostomata (horseshoe crabs):

Development in chelicerata is the less studied among arthropoda. Their eggs are fairly large and yolk rich (except for mites, pycnogonids and scorpions) (Schwager et al., 2015). They normally undergo meroblastic cleavage (most Arachnida (araneomorpha) and Xiphosura) (Kimble et al. 2002)), but both holoblastic (discoidal) (pycnogonids, scorpions, pseudoscorpions and most mites) and mixed (in some araneae, individualized blastomeres appear only after 16 cell stage (Suzuki and Kondo, 1995)) types of cleavage can be observed. It has been argued that holoblastic cleavage might be ancestral mode for chelicerata (actually, some meroblastic cleavages reported in some taxa (e.g. mites) has been proven to be microscopic artifacts: high resolution microscopy has revealed it to be actually holoblastic (Laumann et al., 2010)). However, in pycnogonida, which are supposed to

be the extant earliest divergent taxa among chelicerates, cleavage is almost holoblastic, without traces of yolk (Arango and Wheeler, 2007).

In both mero- and holoblastic cleavers the blastomere arrangement is usually irregular and variable even within a species, retaining no traces of spirality (Ungerer and Scholtz, 2009). In pseudoscorpions, cleavage has been described even as chaotic (Grassé, 1949). However, some transient squared blastomere configurations have been described in the discoidal cleavage of some scorpions. Cell divisions are symmetric in some species and asymmetric in others, like in some pycnogonids (genera *Callipalene* and *Propallene*) (Scholtz and Wolff, 2013; Schwager et al. 2015). It has been argued that this early asymmetric cell divisions are related to a determinate cell specification, although in most chelicerates cell fates are not specified until later development. More cell lineage studies are required to clarify this question (Kanayama et al., 2010).The gastrulation process in chelicerates follow a short to intermediate germ-band development.

3) Crustacea, including shrimps, lobsters, crabs, barnacles and relatives:

Several degrees of completedness in the cleavage furrows are found within crustacea (but normally closer to holoblastic than to meroblastic) (Müller et al., 2004). They have the more regular cleavage pattern among arthropods (respect of its blastomere arrangement, relative sizes and fates), but still great variation in them can be observed (Alwes and Scholtz, 2004). Among the class malacostraca (at least in the orders Amphipoda, euphausiacea and decapoda), the 4-cell stage is akin to the 4-cell stage spiral pattern (see next section), but afterwards the direction of cell divisions is not oblique but perpendicular respect to the AV-axis (Hertzler and Wallis, 1992; Alwes and Scholtz, 2004). Moreover, in amphipoda, the asymmetries in cell size in the 4-cell stage are translated to the next stages, a feature nor found in spiralia (and that probably implies different cellular events acting in the asymmetric cell division) (Wolff and Scholtz, 2002).

In non-malacostracean species, cleavage closely resembles the canonical spiral pattern (Scholtz et al., 2009). This not only includes the obliqueness of cell divisions respect to the AV axis, but also the left-right alternation of these cell divisions after the 4-cell stage (Fig. 2E-F).

However, the cell processes involved in the direction of cell division (cell contacts (Alwes, 2004); Sachs's rule (Pawlak et al., 2010)) the cell fate map and the mode of cell fate determiantion are quite different from those observed in spiralians (Pawlak et al., 2010). Because of that, this “non-spiralian” spiral cleavage has been called modified-spiral cleavage pattern (Valentine, 1977).

In crustaceans, cell fate determination is usually determinative (the cell lineage plays a major role, which differs from the rest of arthropods), and follow at least two different modalities (Price and Patel, 2008): In cirripedia (barnacles), the center of the vegetal pole (future posterior

body end) before gastrulation is occupied by the 4D blastomere (that will give rise to the endoderm, not to the mesoderm as in spiralians) (Valentine, 1997). In cladocerans (*Daphnia*), D-quadrant forms a primordial germ cell and a primordial endoderm cell, but in this case it marks the future anterior body end (not the posterior end, as in cirripedia and spiralians). Moreover, in this species the relative position of blastomeres can be experimentally changed without major consequences for latter development (Kaudewitz, 1950). This suggests that the D cells in true spiralians and crustacea are not homologous (neither their modes of cell-fate determination are) (Siewig, 1979). Rather, the D blastomere of spiralians and crustaceans are labeled the same way simply because of their larger size and the fact they contain most part of yolk (Valentine, 1997; Scholtz and Wolff, 2015). Thus, the modified spiral cleavage pattern has likely evolved independently from the other spiralian cleavage patterns (Siewig, 1979). In general, they exhibit short germ-band or intermediate germ-band type of development.

4) Myriapoda, including centipedes, millipedes and relatives:

Within myriapoda, the classes Pauropoda, Symphyla and Diplopoda feature holoblastic cleavage, whereas most representatives of the class Chilopoda have meroblastic (intralecithal) cleavage. Within the forms displaying intralecithal cleavage, some taxa have the yolk partitioned in irregular ways while some others (Pleurostigmophoran centipedes) have it partitioned in regular sectors that depart from the embryo center, where the cells´ nucleus are located (a cross section of this blastula is a radially symmetric rosette). In these special case (intrapyramidal cleavage), when the central cells have divided a number of times, they migrate radially to the embryo´s surface between the yolk compartments to form the blastoderm (Brena, 2015).

Overall, the data do not allow to infer unambiguously the ancestral cleavage pattern for Miriapoda (Scholtz and Wolff, 2013). In the cases where cleavage is fully holoblastic (eg. The millipede *Glomeris*), the first two cell divisions are symmetric and meridional (giving rise to a four equal blastomeres that lie in the same plane). After this 4-cell stage, both symmetry and synchrony are progressively lost in the ensuing cell divisions, and not clear geometrical pattern can be distinguished in the blastomere arrangement (Scholtz and Wolff, 2013). As a general rule, myriapods follow a short or intermediate germ-band formation (Chipman at al., 2004).

3.3.5. Onychophora:

Onychophora (aka velvet worms) is a small phylum related to arthropoda (Arthopoda, Onychophora, and tardigrada (see below) form together an unranked clade called Panarthropoda).

They are millipede-shaped animals with an obscurely segmented body covered by a non chitinized cuticula. Several reproductive strategies have been described, with oviparous, ovo-viviparous and viviparous forms. The longer the external (out of the mother) development the bigger and yolkier the eggs, and the more resistant and protective the chitinous eggshell (composed of several specialized envelopes) (Mayer et al., 2015). Thus, in viviparous species cleavage is meroblastic (discoidal) and finishes in a monolayered hollow blastula (Anderson and Manton,1972). Oviparous species, in turn, tend to display holoblastic cleavage (Anderson, 1966; Eriksson and Tait, 2012).

In the species exhibiting discoidal cleavage, first cell divisions tend to occur at right angles to the previous cell division (Sachs' rule), in a synchronous manner. That way, blastomeres in 8-cell stage form two parallel rows of four cells each, and by 16-cell stage a four by four square of equally sized cells can be seen in one side of the elongated embryo. Both the squared arrangement of cells and synchronization are lost after 32-cell stage, when the cellularization of the energids takes place (Eriksson & Tait, 2012; Meyer et al., 2015).

The blastomeres continue dividing and spreading to eventually cover the surface of the egg, forming a blastoderm. The blastoderm covers the central yolk mass, which split into separate, rounded yolk compartments after 64-cell stage (Eriksson and Tait, 2012). During gastrulation, only endoderm pass through the blastopore, which closes after that. Subsequents events are more variable between species (Eriksson et al., 2012; Meyer et al., 2015).

Some aspects of the onychophoran cell fate map (e.g. the mesoderm is formed behind the endoderm) show similarities with the annelids, which for some authors (Anderson, 1973) points to a deep relationship between them.

3.3.6. Tardigrada

Tardigrada, or “water bears”, are microscopic animals of marine and terrestrial soils which can form resistant (encysted) forms that can survive to very extreme environmental conditions (Rupert et al., 2004). Eggs (normally sub-spherical) are laid either directly into environment or into the old exuvium just after ecdysis (Brusca and Brusca, 1990). The first ones have external motifs of taxonomical interest, while the second ones are generally smooth.

Cleavage is total (holoblastic) with evenly distributed yolk droplets (Scholtz and Wolff, 2013). Between species, different cleavage patterns have been reported, ranking from regular to irregular, from initially synchronous to always asynchronous and from symmetric to asymmetric after the 4-cell stage (Gross et al., 2015). In the emerging model species *Hypsibius dujardini* (smooth egg, stereotyped, initially synchronous and symmetric), first two cell divisions are

meridional. In the ensuing cell division, the two cells divide asymmetrically (and the smaller cells continue dividing at slower pace until the 60-cell stage, when a compact stereoblastula is formed and gastrulation begins (Gross et al., 2015). Gastrulation is performed via directed cell migration to the more central regions of the embryo, being the progeny of the asymmetrically dividing cells the first ones that enter. Nor the mechanisms of body-axis specification nor the cell fate map have been adequately described for Tardigrades.

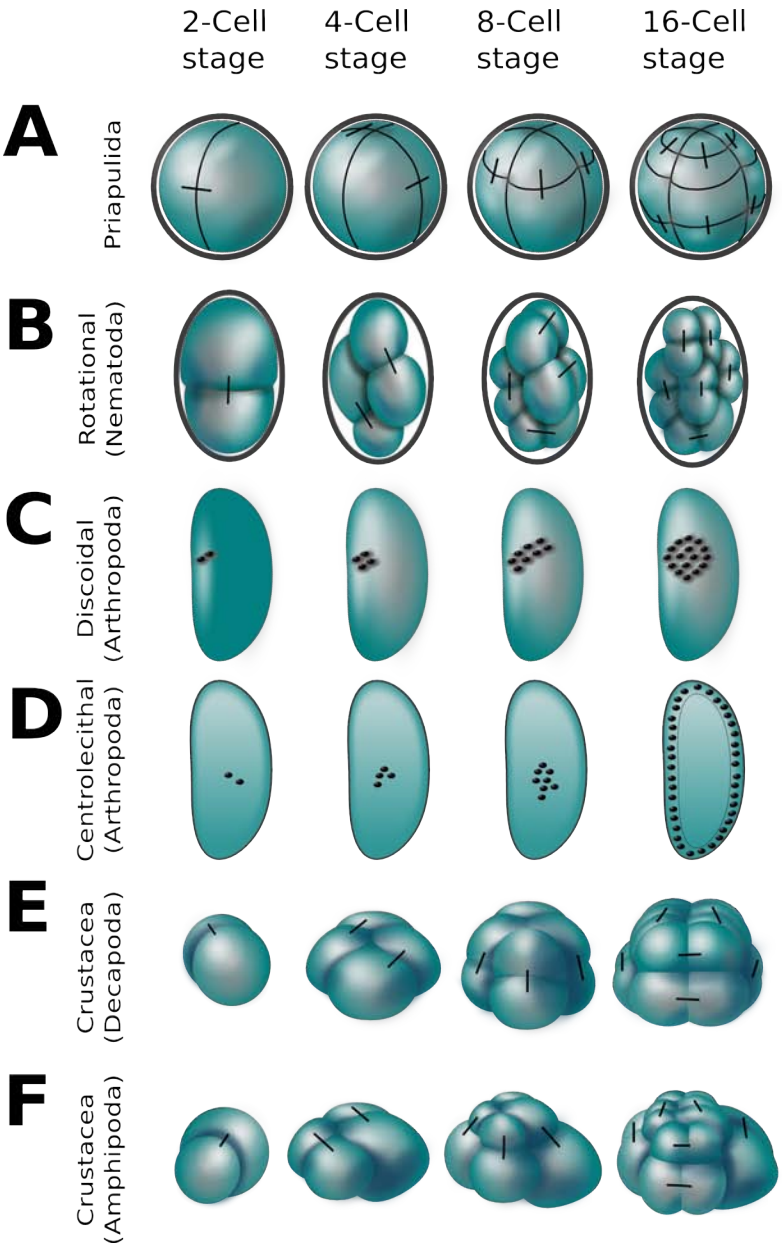


Figure 2 (From the previous page). Example cleavage patterns found in the superphylum *Ecdysozoa*. Most of them may be explained by means of the combination of a few conserved processes (see text). All blastulae are displayed in lateral view with the animal pole on the top, and the small straight lines link sister blastomeres when both of them are visible and identifiable. A) The cleavage of Priapulida, probably driven by a Sachs-like rule. B) Rotational cleavage pattern of *C. elegans* (a model species representative of the phylum Nematoda). C) Discoidal cleavage of some Arthropodan taxa D) Centrolecithal cleavage of some Arthropodan taxa (pictures show a longitudinal section of the embryo, and the right-most picture is not a 16 cell-stage but the stage of the blastoderm formation (see text)). E) Holoblastic cleavage of Decaphoda (crustacea). F) Holoblastic cleavage of Amphipoda (crustacea).

3.4. Spiralia.

Spiralia (aka Lophotrochozoa) is a very diverse group of animals that comprises almost half of the animal phyla. Molecular evidence from the 18S ribosomal RNA suggests that they are a monophyletic group (Halanych et al., 1995; Hervé et al., 2005;). Their monophyly is also supported by the fact that most of them share several morphological traits, such as the presence of ciliated (trochophore or trochophore-like) larvae and the arrangement of the blastomeres in their early cleavage (Nielsen, 1994; Laumer et al., 2015). Most of the spiralian phyla (Mollusca, Annelida, Nemertea, Platyhelminthes, Entoprocta and Gnathostomulida), in spite of having very different adult morphologies, exhibit a very conserved cleavage pattern called spiral, “canonical spiral” or “quartet spiral” (Freeman and Lundelius, 1992; Hejnol, 2010; Henry, 2014). Some deviations of this pattern do occur in some spiralian groups. Such deviations are spiral-like patterns in which one or several features of the quartet spiral pattern (a complex trait itself) have been lost or modified (Jenner, 2004; Hejnol, 2010). Let us first describe the main features of the quartet spiral pattern, and then its main evolutionary modifications.

3.4.1. The quartet spiral pattern:

Most of spiralians develop into a planktotrophic free-swimming trochophora larva capable of self-sustaining, so that yolk is no longer required for nourishment and consequently the cleavage is holoblastic. The quartet spiralian cleavage begins with two meridional cell divisions giving rise to four large macromeres, called A,B,C, and D macromeres. In the quartet spiral pattern (and also in some spiral-like patterns (Henry et al., 2000; Hejnol, 2015)), the regularity of the blastomere arrangement allows the labeling of individual blastomeres in a standardized manner. This labeling was developed in 1897 by the embryologist E.G. Conklin, and follows these rules (Conklin, 1897; Gilbert and Raunio, 1997):

a) macromeres in the vegetal pole are defined by a capital letter corresponding to their quadrant preceded by the number of cell divisions that have happened after the 4-cell stage.

b) the sister cell of each of the four vegetal-most blastomeres are labeled the same way, but with low-case letters,

c) the label of each one of the remaining blastomeres are the same as its mother cell with a superindex after it: 1 for the daughter cell closer to the animal pole, and 2 for the daughter cell closer to the vegetal pole.

The cell lineage of each of those is called a “quadrant” (e.g. D-quadrant), because when viewed from the animal pole, each cell lineage occupies a sector of about 90 degrees of the circular contour of the blastulae (that is, a geometrical quadrant). These macromeres then divide towards the animal pole but at an oblique angle relative to the animal-vegetal axis, giving rise to four, normally smaller animal micromeres that are all displaced to the right (or all to the left depending on the organism) of its sister macromere (see Fig. 3A).

This relative tilt or handedness (micromeres to the right or to the left of their sister macromeres) during this 8-cell stage has deep consequences for later development (for instance, snails with a tilt to the right have a dextrally coiled shell) (Grande and Patel, 2008; Kuroda et al., 2009), and it is not randomly determined (that is, in a given species the proportion of right and left tilts is not 50%-50%). Crossing experiments show that the handedness is strongly determined by a maternally inherited cytoplasmatic factor(s), and in most spiralian species, wild type populations have the same handedness, being thus all adults individuals either left or right-coilers (since mixed populations would be counteradaptive, adult enantiomorphs (left and right coilers) are sexually incompatible between them) (Freeman and Lundelius, 1982; Asami et al., 1998). This genetic determinaation can be overcomed by mechanical manipulation of the 8-cell stage embryos (a pushing of each micromere to its adjacent furrow between macromeres) (Kuroda et al., 2009).

There are a number of developmental processes that have been hypothesized to explain how this tilt between micro- and macromeres arises in the 8-cell stage of the spiral cleavage. Some authors propose that complex short-range signaling can orient the mitotic spindles in a coordinated manner prior to cell division (Freeman and Lundelius, 1972; Aw & Levin, 2009). Other authors propose that cortical rotation after cell division is the main driver of this symmetry breaking (Meshcetryakov and Beloussov, 1975). Others consider that several mechanisms may synergistically contribute to the emergence of spiral pattern (Kuroda et al., 2009; Shibazaki et al., 2004; Wandelt & Nagy, 2004). For instance, in *Lymnaea* gastropods the dextral (wild type) and

sinistral tilts between micro- and macromeres are not attained by equivalent (but opposed) cellular processes. Instead, the dextral snails exhibit additional cell features not observed during cleavage in sinistrals (spiral deformation SD and spindle inclination SI). Since an effective sinistral tilt is realized, it follows that neither the SD nor the SI are required, and that a simpler way to build an spiral pattern does exist (Wandelt & Nagy, 2004). This point of view has been recently assessed by means of a computational model of spiralian cleavage that combines all cell processes known to be involved in early cleavage (Brun-Usan et al., in press), This work shows that a combination of cortical rotation and directed mitosis towards the animal pole (plus cell adhesion) are necessary and sufficient conditions for the attainment of the 8-cell stage spiral (ibid).

After the 8-cell stage, the ensuing cell divisions follow a right-left alternation (the reverse alternation applies if the third division is to the left), making that, when viewed from the animal pole, the new micromeres seem to spin clockwise or counterclockwise when they arise. Guerrier (1970) suggested that this alternation between clockwise and couterclockwise rounds of cell division may be explained by assuming that each cell division is perpendicular to the previous one (Sachs' rule) (Minc & Piel, 2012). However, this explanation has two caveats : 1) There are infinite directions (albeit in the same plane) perpendicular to a previous cell division. 2) The initial symmetry breaking tilt between 4 and 8 cell stage remains unexplained. Thus, as some author suggest, additional cell processes, such as directed cell division towards the animal pole of the embryo, should operate along with Sachs' rule in order to produce this left-right alternation beyond the 8-cell stage (Brun-Usan et al., *in press*).

As the quartet spiral cleavage proceeds, each set of four blastomeres that, belonging each one to a different quadrant, are located in the same relative position along the AV axis (because they were released simultaneously) is called a “quartet” (e.g. the blastomeres $2a^1$, $2b^1$, $2c^1$ and $2d^1$ form a quartet).

When cell fates are compared between different spiralians, the same adult or larval organs in different species arise from the same blastomeres (defined by lineage and relative position in the blastula). However, the mode of cell fate determination slightly differs between quartet spiralian cleavers. Specifically, the so called “D-blastomere” (a mesodermal precursor) can be specified by three different mechanism (Freeman and Lundelius, 1992):

First, by a contact mediated inductive interaction between one of the macromeres at the vegetal pole with micromeres at the animal pole of the embryo after the fifth cell division (equal spiral cleavage) (Biggelaar and Guerrier, 1983; Lambert et al., 2003; Grande & Patel, 2008; Aw & Levin, 2009)

Second, by asymmetrical segregation of cytoplasmic determinants, which in turn is caused by asymmetric cell division the 4-cell stage (Fig. 3B). In this case, that is called unequal spiral cleavage, the D-blastomere remains visible larger than the other ones (A,B,C) in the four cell-stage. In some cases (e.g. the freshwater clams of the family *Unionidae*) the asymmetric division of the D-blastomere lineage continues until the 8-cell stage. The resulting 2d-blastomere, that is much larger than the 2D-blastomere inherits most of the cytoplasm of it immediate precursor,the D-blastomere (this probably implies a mechanism other than animal-vegetal gradient driving this asymmetric cell division). This 2d blastomere is the precursor of specialized shell-like larval structures, by which the *Unionidae* larva (*glochidium*) is dispersed by getting fixed to the fishes' gills.

The third mechanism involves the formation of an big and transient cell protrusion, the polar lobe, at the vegetal pole during cytokinesis. The polar lobe engulfs a substantial volume of the blastomere's cytoplasm and cytoplasmic determinants during a critic inductive period, causing an asymmetry in the 4-cell stage and the specification of the D-lineage (the removal of the polar lobe leads to radialized embryos in which any blastomere is specified as D).

Phylogenetic analysis suggest that the first mechanism of D-quadrant specification (equal spiral cleavage, inductive determination at later developmental stages) is ancestral in spiralians, and the others have arisen several times in different lineages (Freeman and Lundelius, 1992).

Summarizing, the spiralian development, in general, is highly determinative (operating within a constant blastomere arrangement), with some critical points driven by inductive signaling events between adjacent blastomeres (e.g. the choice of the embryo's handedness or the lying out of the D-blastomere) (Freeman and Lundelius, 1982; Kuroda et al., 2009; Grande and Patel, 2008; Henry, 2014). In most cases, gastrulation occurs through epiboly of the animal micromeres over the vegetal cells (Gilbert and Raunio, 1997).

3.4.2. Evolutionary modifications of the spiral pattern:

So far, we have reviewed the spiralian cleavage patterns whose features are compatible with the quartet spiral pattern, but in some Spiralians, the quartet spiral pattern has been total or partially lost by different causes. In some cases, they represent exception within phyla whose members undergo quartet spiral cleavage. For instance, massive amount of yolk correlates with the lost of the spiral pattern (and a switch to a specific meroblastic cleavage) in cephalopod mollusks (Gilbert and Raunio, 1997) (Fig. 3C). Another example is provided by leeches (clitellate annelids), in which some cytoplasmic animal and vegetal domains (which are yolk deficient but enriched in

mitochondria and maternal mRNAs) proliferate and form a chain of small cells (teloblasts) that will give rise to the segmental tissues of the annelids, disrupting the quartet spiral pattern after the 8-cell stage (Holton et al., 1994; Weisblat and Huang, 2001)

In other cases, almost all representatives of the phylum have lost the quartet spiral pattern:

3.4.2.1. Rotifera:

The rotifera (or wheel animals) is a phylum of microscopic planktonic animals which bear an anterior ciliary ring, called corona, devoted to feeding and locomotion. Adults frequently exhibit eutely, meaning that they have a fixed number (about one thousand) of cells within a species, that are generated early in development (Rupert and Barnes, 1994). Contrary to other spiralians, left-right asymmetries in the adult body are restricted to the position of some ovary structures (Namigai et al., 2014). Rotifera comprises two broad classes: *bdelloidea*, with telescopic cylindrical body and usually divided corona; and *monogononta*, with one ovary and reduced corona (Brusca and Brusca, 1990). *Seisonidea*, a single genus comensal of crustaceans is often considered as a third class, but its phylogenetic relationships, including its placement within rotifera, are still debated (Zrzavy, 2013; Sielaff et al. 2015). Larval stages only exist in sessile species (Rupert and Barnes, 1994).
Most rotiferan groups are parthenogenetic, but dwarf males also do exist within *monogononta* (Brusca and Brusca, 1990). Inside the females, yolk is accumulated around the oocyte nucleus before being pinched off from the syncitial ovary as a mature oval egg, which is surrounded by a number of shells secreted by the egg itself (Rupert and Barnes, 1994).

Development in the class monogononta has been more extensively studied than in bdelloidea, but the early development does not differ dramatically between the two classes (Hejnol, 2015). As in other spiralians, cleavage in rotifera is determinate (Rupert and Barnes, 1994; Nielsen, 2005).

Just before cleavage, one polar body is extruded, defining the animal pole of the oval embryo. First cell division is asymmetric, and produces an animal (AB, ectoderm precursor) blastomere and a larger vegetal (CD) one. In the second cell division round, AB blastomere divides symmetrically into A and B blastomeres, in a direction perpendicular to the AV axis (and to the previous cell division). In turn, CD blastomere divides roughly in the same direction, but asymmetrically, into the C blastomere, which occupies the medium region of the embryo, and the much smaller D blastomere (see Fig. 3D).

The resulting configuration of the four-cell stage looks similar to the early development of some unequal cleaving spiralians: three roughly similar blastomeres (A,B, and C) that are adjacent

to a much larger D blastomere, and the polar body in the center of these four blastomeres (Wallace et al., 1996; Hejnol, 2015) (see section 3.4.1). The two subsequent cell divisions rounds are oriented along AV axis, and except for the D blastomere, are roughly symmetric. Contrary to other spiralians, the cell lineage of each one of the four initial blastomeres do not follow a right-left alternation respect to the AV axis, but forms straight columns of compressed blastomeres along it. Moreover, the overall shape of the the embryo tracks the shape of the ellipsoidal eggshell, suggesting that the compressive effect of eggshell may have a role in the blastomere arrangement after the 4-cell stage (e.g. by hindering the relative twist between micro- and macromeres due to cortical rotation) (Schierenberg and Junkersdorf, 1992; Kajita et al., 2003; Wanniger, 2015).

After the 8-cell stage, the polar body is displaced to the vegetal pole in monogononta (Hejnol, 2015). Despite this displacement, the fate of each blastomere is comparable between both rotiferan taxa, suggesting that cell fates are determined early and that no inductive mechanisms between neighboring cells are required for that (Lechner, 1966; Hejnol, 2015). Inductive mechanisms in later development, however, seem to play a role in the specification of the germ line, at least in some rotiferan genera (e.g. *Brachionus*) (Smith et al., 2010).

When the embryo reaches the 16-cell stage (some bdelloid genera, such as *Callidina* or *Macrotrachela*, pass through transient 10-cell stage), gastrulation begins by the internalization of the D blastomere, which remains much larger than the others (Zelinka, 1892; Boschetti et al., 2005). This D blastomere divides while internalizing and, after that, other cells follow the internalization process.

In rotifers, D blastomere will give rise chiefly to the ventral side of the body, while in other spiralians it represents the dorsal side. Because of all these differences in both blastomere arrangement and cell fate maps, the interpretation of the rotiferan cleavage as “modified” spiral is not clear, and appears to hinge upon the phylogenetic position of Rotifera among the other spiral cleaving phyla (Costello and Henley, 1976; Valentine, 1997; Jenner, 2004; Zrzavy, 2013).

3.4.2.2. Acanthocephala:

Acantocephalans are parasitic worms characterized by the presence of a protrusible spiny proboscis, and comprises about 1150 species (Brusca and Brusca, 1990). Most of phylogenetic studies place them as the closest taxon to rotiferans (even they have been described as a highly modified rotiferans), forming together a taxon called Syndermata (Garcia-Varela and Nadler, 2006; Sielaff, 2015). However, their early development, which has been studied in few species and seems to be rather homogeneous between them, is not very similar to this observed in rotiferans (Schmidt,

1985; Hejnol, 2015). Their cell fate map has not been defined unambiguously, but early development is known to finish in the *Acanthor* larva, the first of several infective stages that characterizes the complex life cycle of Acanthocephalans.

In general (but the events described in here correspond to the genus *Giganthorhynchus*), eggs are extremely elongated, and are surrounded by a four-layered system of hard membranes (Uglem and Larson, 1969). Two polar bodies mark the anterior (animal) pole of the embryo in one of the egg's apexes (Meyer, 1928; Schmidt, 1985; Hejnol, 2015). First cell division (along AP axis, the major axis of the embryo) is asymmetric, being the resulting anterior (animal, AB) blastomere slightly smaller than the posterior (vegetal, CD) blastomere (see Fig. 3E). In the second cell division round, both blastomeres divide again in the same direction in a symmetric manner. In the four-cell stage, there are one blastomere (B3) in the animal pole, one (D3) in the vegetal pole and two “equatorial” blastomeres (A3 and C3) in the medium region of the embryo. The nucleus of the A3 blastomere (descendant of the animal blastomere AB) is located in a more vegetal position than the A3 blastomere (descendant of vegetal blastomere CD) (Meyer, 1928; Hejnol, 2015). Whether this spatial blastomere configuration results from the packing constraints imposed by the elongated eggshell, from some movement performed by blastomeres after cell division or by directed cell division is not clear. In the next cell division round all blastomeres divide asymmetrically (approximately along AP axis), each one giving rise to a micromere (closer to the animal pole) and a macromere (closer to the vegetal pole), thus generating a unique cleavage pattern that has a sort of bilateral symmetry around the AP axis.

In the the next (fourth) cell division round, the division or micromeres are strongly delayed or even arrested, making further cleavage asynchronous. Moreover, in different species the embryo becomes syncitial before gastrulation, which proceed by internalizing the micromeres in a “central nuclear mass”. Thus, Acanthocephlans exhibit what some authors term as “distorted” spiral cleavage pattern. However, no features of the quartet spiral pattern are found (Crompton, 1989; Nielsen, 2001; Jenner, 2004) because from the very beginning, the A,B,C and D cell lineages never occupy the same relative position along the AV axis (as they do in form of quartets in the canonical spiral cleavers and rotiferans). It is tempting to speculate that this may be due to the dramatic spatial restrictions derived from the very elongated and stiff eggshell, that in turn represent an adaptation to their parasitic way of life (Uglem and Larson, 1969), but if this is actually the case has to be checked (Wanniger, 2015).

3.4.2.3. Gnathostomulida + Micrognathozoa:

Ganthostomulids are a small phylum of nearly microscopic, marine interstitial worms characterized by a monociliary epidermis and a blind gut with a jaw-like structure on its opening (Brusca and Brusca, 1990; Rupert and Barnes, 1994; Ax, 2003). Their early development is only known for one species: *Gnathostomula jenneri* (Riedl, 1969; Hejnol, 2015). Its early development up to the 8-cell stage is akin to the quartet (equal) spiral cleavers, including the presence of four animal micromeres with a dexiotropic spiral tilt (Fig. 3F). However, after the next cell division (in the 16-cell stage), all but the four larger vegetal-most blastomeres are arranged in the same layer, close to the animal pole, in a nearly squared configuration, that resembles (but is not identical to) the "annelid cross" pattern (Riedl, 1969), although as explained above, the phyogenetic status of this feature is no longer tenable (Jenner, 2004; Maslakova et al., 2004). This flattening of the embryo can be due either to a differential adhesion between blastomeres (low between micromeres, high between micro- and macromeres) or to a weakening of the polarizing effect of the animal pole in orienting cell divisions.

Nothing is known about the early development of *Lymnognathia maerski*, the only described species of the closely related phyla micrognathozoa (Kristensen and Funch, 2000).

3.4.2.4. Gastrotricha:

Gastrotricha, also know as hairybacks, is a phylum of tiny marine worm-like animals whose early development takes place inside an ellipsoidal and often spiny eggshell that is not maternally provided but is produced by the egg itself (Rieger and Rieger, 1980). The polar bodies are extruded just after egg deposition to one tip of the ellipsoidal embryo, defining the animal pole (which correlates with the anterior part of the juvenile), but the two classes that compose this phylum exhibit different cleavage patterns (Sacks, 1955; Hejnol, 2015).

In the class Chaetonotida (characterized by the absence of pores in the pharynx, and by the presence of posterior adhesive glands), the first cell division is equatorial and symmetric, and gives rise to an animal (AB) and a vegetal (CD) blastomeres (see Fig. 3H) (de Beauchamp, 1929; Sacks, 1955; Rupert and Barnes, 1994; Hejnol, 2015). Then, the AB blastomere divides into the A (larger) and B (smaller) blastomeres; and just after that the CD blastomere divides into the equally sized C and D blastomeres (Hejnol, 2015). These two latter twist something less than 90 degrees around the AP axis just after cell division, being now the pairs of animal and vegetal blastomeres oblique one to another. In the third division round, all cell divisions are roughly symmetric, and blastomeres

divide perpendicularly to both the AV axis and to the previous cell division (Hejnol, 2015). The resulting 8-cell stage configuration is reminiscent of the quartet spiral pattern (or even the duet spiral one, see section 3.2), but in this case the blastula is strongly elongated along AV axis and two bigger blastomeres (b.3.2 and a.3.2) are located near the animal pole. In the next cell division round, all blastomeres divide along the AP axis (Sacks, 1955).

In the less studied class Macrodasyoida, characterized by the presence of two pores (involved in osmotic regulation) on either side of the pharynx, the events until the four-cell stage are rather similar. However, the relative twist of animal and vegetal blastomeres in the 4-cell stage is more pronounced than in the class chaetonotida (about 90 degrees), because vegetal blastomeres divide perpendicularly to the animal ones, so that they attain an alternate configuration without additional rotation (Teuchert, 1968). Then, in the 4-cell stage, the vegetal C-blastomere moves towards the animal pole, which leaves a single (D) vegetal blastomere (ectoderm precursor, Teuchert, 1968) and three animal ones (Fig. 3G). After that, all blastomeres divide along the AV axis, but the C-blastomere's progeny is located in a much more vegetal position than the descendants of the other (A and B) animal blastomeres. In the next cell division, cells divide perpendicularly to the previous cell division (and to the AV axis).

During cleavage, in all Gastroctrichs the blastomeres close to the animal pole always divide slightly earlier than the ones close to the vegetal pole, but cellular mechanisms involved in determining the direction of cell divisions and the placement of the blastomeres are not clear (the perpendicularity between most of consecutive cell divisions, however, points to a Sachs-like rule, although an active role of the elongated eggshell in determining the blastomere arrangement can not be discarded).

In both classes of gastrotricha, gastrulation begins around fifth cell division (32-cell stage in Chaetonotida and 30-cell stage in Macrodasyoida) by the migration of two laterally-located blastomeres (descendants of the A-blastomere in chaetonotida, and of the D-blastomere in Macrodasyoida) to the blastocoel, which is visible in the center of the blastula (Sacks, 1955; Teuchert, 1968). These early internalized blastomeres are endoderm precursors, and are supposed to be homologous between classes, although the detailed cell fate maps of the two classes have not been adequately established (Teuchert, 1968; Hejnol, 2015).

3.4.2.5. Platyhelminthes:

Platyhelminthes (flatworms, tapeworms) is a phylum of small, unsegmented and often parasitic bilaterian animals, capable of asexual (budding) and sexual reproduction (Cardona et al.,

2005). Their relatively simple organization (they show no body cavity, specialized circulatory or respiratory organs) has challenged taxonomists, and traditionally, Platyhelminthes were classified according to their lifestyles in Turbellaria (generally free-living flatworms) and Neodermata (parasitic forms). However, in 1985, Ehlers proved that Turbellaria was paraphyletic, after which dozens of evolutionary schemas have been proposed for this phylum (Ehlers, 1985; Timothy et al., 2004; Egger et al., 2015). However, the term Turbellaria, *sensu lato*, still continues being used (Martin-Duran and Egger, 2012).

When they reproduce sexually, most of the free-living forms (most Turbellaria) spawn a moderate amount of eggs with few yolk, and exhibit a quartet spiral cleavage pattern, including the cell-fate map (Boyer et al., 1998, Yunossi-Hartenstein and Hartenstein, 2000; Rawlinson, 2010; Lapraz et al., 2013). On the other hand, the parasitic forms (a few parasitic Turbellarian forms and Neodermata), which often have complex life cycles, spawn thousands of tiny, yolky eggs that lack all traces of spiral cleavage (Martin-Duran and Egger, 2012). In the latter, eggs have several protective membranes (the so called egg-capsule), and generally show an opening (operculum) for larval hatching. The egg-capsule is secreted partially by the Mehlis' gland, a multi-purpose secretive organ adjacent to the ootype, and partially by the egg itself. An additional protective structure called hull membrane, which consists on many flattened and sometimes syncitial cells (blastomere derivatives) lining the egg surface is often found in both free-living and parasitic groups (e.g. Macrostomida, Aspidogastrea) (Swiederski et al., 2011).

In general (but more often in parasitic groups), the oocyte, that is much smaller than the egg-capsule, is surrounded by several yolk-filled nurse cells (vitellocytes) which were provided by the vitellarium (an ovary accessory gland). Both vitellocytes and the central oocyte are located inside the egg-capsule, a configuration known as Ectolecithal cleavage (Rupert and Barnes, 1994). Furthermore, in some groups a number (e.g. three in the case of Tricladida) of whole eggs, each one with its corresponding oocyte and yolk-cells, are located inside a larger structure called egg-capsule, that contains a large population of extra embryonic yolk-cells. (Stevens, 1904; Cardona et al., 2006).

The purpose of these vitellocytes is twofold: they provide the nutrition of the embryo and they are recruited for the building of specialized protective structures around the embryo: parts of the inner egg-shell and, in Rhabdocoela and Tricladida also the hull membrane (Martin-Duran and Egger, 2012). In many species, vitellocytes fuse into a perypheral syncytium during early cleavage (Cardona et al., 2006). The large number of yolk-cells in the Platyhelminthes egg is thought to be an adaptation to optimize the rate of yolk production. Since the small molecules used for yolk synthesis have to traverse the cell membrane, the total membrane area of the yolk-supplier cells

appears as a limiting factor. This may be partially overcomed by splitting a single large yolk-cell in may smaller ones (for a given volume, smaller cells have larger area than a single big one), so that the total area membrane increases and the rate of yolk synthesis speeds up (Rupert and Barnes, 1994).

In blastomeres, cell divisions are initially symmetric and synchronous, but both features are quickly lost. In parasitic forms, the greater the presence of yolk-cells the more the cleavage pattern departs from the spiral one (some authors directly attribute the disturbance of the spiral pattern to the pressure exerted by the surrounding yolk-cells (Newton, 1970; Rupert and Barnes, 1994)). As a general rule, in the forms that do not follow the spiral cleavage pattern the cleavage pattern can not be exactly established, but it seems to be irregular (chaotic) (Willey and Goodman, 1951; Swiderski et al., 2011). Even in some taxa (Tricladida, bothrioplanida and prolecitophora) detached blastomeres float adrift in the bulk of yolk, a condition referred to as disperse cleavage or *Blastomerenanarchie* (Martin-Duran and Egger, 2012) (see Fig. 3I). Some blastomeres and vitellocytes, after having contributed to the the eggshell formation and the embryo nourishment, degenerate and undergo apoptosis (Swiderski et al., 2011).

The topological relationship between the vitellocytes and the mass of blastomeres vary as development proceeds. At the beginning, blastomeres usually occupy the centre of the blastula, and then they are displaced to the periphery to form a pseudoepithelial sheet of cells (blastema). In some groups (Lecithopitheliata, Proseriata, Bothrioplanida and Rhabdocoela) the mass of yolk cells are only partially covered by the blastema, whereas in others (Fecampiida, Prolecithophora and Tricladida), the layer of blastomeres eventually enclose all (or a substantial portion of) the yolk mass. In the case of the Tricladida, the engulfment process is rather special (exclusive for tricladida) and deserves a description on its own: a subpopulation of the dispersed blastomeres form a transient (embryonic) epidermis that surround the yolk syncytium (as in some of the already mentioned groups) and a transient (embryonic) pharynx (a sort of cylinder made of a handful of cells located in one side of the embryo). Then, extraembryonic yolk cells are actively swallowed into a cavity that opens in the center of the embryo. After the yolk ingestion, the remaining subpopulation of undifferentiated blastomeres (those which did not participate in the formation of the embryonic epidermis nor the embryonic pharynx) proliferate and differentiate into the definitive larval organs, that replace the transient ones (thus, the embryonic point where the transient embryonic pharynx appears does not correlate with the placement of the adult definite pharynx or gut).

Most of the groups show gastrulation by epiboly (the expansion of a sheet of blastodermal cells over the others, so that the “external” blastoderm lines itself by creating an internal sheet of cells). In the forms which conserve the spiral cleavage, the animal micromeres usually expand over

the vegetal macromeres (stereoblastula) during epibolic gastrulation). In the coeloblastula of Macrostomorpha and prolecitophora, the blastoderm (blastema), is initially internal, and then unfolds to create a sheet of external cells, a process called “inverse epiboly” (it is thought not to be homologous to the “normal epiboly” found in other groups). Finally, in Proseriata and Tricladida the gastrulation process is very idiosyncratic, showing features of different processes (Martin-Duran and Egger, 2012).

As a general rule, their development is determinative (early cell-fate specification by unevenly distribution of cytoplasmic factors in the oocyte) in the groups that still conserve the spiral cleavage pattern and regulative (cell-fate determination through cell-cell communication processes in the embryo) in those which do not (Rupert and Barnes, 1994; Martin-Duran and Egger, 2012).

3.4.2.6. Lophophorates

Lophophorates, a superphylum that contains the brachiopoda (lamp shells), bryozoa (aka ectoprocta) and phoronidea (Nielsen, 2012; Laumer et al., 2015; Schmidt-Rhaesa, 2007), are sessile animals which possess a complex food-catching tentacular organ called lophophore (Rupert and Barnes, 1994). They are included within spiralia (Lophotrochozoa hypothesis; Giribet, 2008) on the basis of some developmental features, such as their trochophora-like larvae (Cyphonautes, actinotroch...), as well as by well supported phylogenomic analysis (Strom, 1977; Rupert and Barnes, 1994; Giribet, 2008).

However, most of them have lost the canonical quartet spiral pattern (Lambert, 2010; Henry, 2015). In lophophorates, in general the eggs are spherical (although some Phoronidean eggs have irregular shaped because they are released in form of aggregates), and the polar bodies define the animal pole of the embryo (Malakhov, 2000). The cleavage is always holoblastic, exhibiting very rounded and loosely attached blastomeres (likely due to a weak cell-cell adhesion) that often leaves visible gaps between them and a prominent central blastocoel (Freeman, 1999; Malakhov, 2000; Pennestorfer, 2012). However, in some colonial forms of Ectoptocta that lack dispersive planktotrophic larvae, the eggs (which are brooded inside a special reproductive zooids called gonozooids) are very yolky and the cleavage is almost (but not totally) meroblastic (Strom, 1977; Zimmer and Woollacott, 1977). These colonial forms often produce multiple embryos by fragmentation of the early cleavage stages (polyembriony) (Rupert and Barnes, 1994; Zimmer and Woollacott, 1977).

Contrary to quartet spiral cleavers all cell divisions are symmetric, and after the first meridional cell division, perpendicular to the previous ones (Sachs´ rule), which leads to a radial-

like (cube-shaped) pattern in the 8-cell stage (Conklin, 1902; Rupert and Barnes, 1994; Freeman, 1999; Malakhov, 2000).

Then, the cell divisions continue being perpendicular to each other until the 32 cell-stage (without the oblique tilt found in the quartet spiral pattern), see Fig. 4A. That means that in the 16-cell stage there are four tiers of four blastomeres in line, and in the 32 cell stage there are eight tiers of four blastomeres in line along the AV axis (Rupert and Barnes, 1994; Wood, 2007). This cleavage pattern, characteristic of lophophorates is called biradial (actually, it can be thought as a special case of radial cleavage because each cell division is perpendicular to the previous one and to the AV axis, but in this case in the fourth cell division round the direction of all cell divisions are the same). This biradial pattern is slightly disturbed in different lophophorate taxa (Long and Stricker, 1991; Hejnol, 2010):

In Brachiopoda the biradial pattern becomes irregular mainly due to asynchronous cell divisions that become generalized after 16 cell-stage (*Terebratulina*, Conklin, 1902), 32-cell stage (*Crania*, Nielsen, 1991) or 64 -cell stage (*Glottidia*, *Dincinisca*, Freeman, 1999). Thus, pre-gastrula stage in brachiopoda contains blastomeres of different size (due to a delay in the cell division timing of some cells rather than to asymmetric cell division events) which, in addition, become irregularly shaped by the mechanical compression exerted by the neighboring blastomeres (Conklin, 1902; Pennington et al., 1999).

In colonial ectoprocts, in the vegetal pole of the embryo the cell divisions are asymmetric due to yolk, causing the four central cells of the vegetal pole to be bigger than the others (Strom, 1977; Gruhl, 2010).

Within phoronidea, although most species (*Phoronis ijimai*, *Phoronis architecta* and *Phoronis buskiit*) are known to display biradial cleavage, in others (*Phoronis muelleri* and *Phoronopsis viridis*) the cleavage pattern has been reported to be spiral (Rattenbury, 1954; Pennestorfer and Scholtz, 2012) or even a mixture of both radial and spiral (Emig, 1974), see Fig. 4B. However, in the reports of spiral patterns there exists great intraespecific variability in the positioning of the blastomeres (in the 16-cell stage one third of the analyzed embryos of *P. muelleri* display radial features), and this variability increases with developmental time (Rattenbury, 1954; Malakhov, 2000; Pennestorfer, 2012).

These intra-specific coexistence of radial-like and spiral pattern in some lophophorates, if is not observational artifacts (Emig, 1974; Brusca and Brusca, 1990), suggests that, at least within this group, small differences in cell mechanics (e.g. cell adhesion) can lead to drastic effects in the resulting blastula configuration (as suggested by Freeman, 1999). Greater differences in cell mechanics may account for the modification, simplification or loss of the spiral cleavage pattern in

lophophorates and other spiralians (Henry and Martindale, 1999; Hejnol, 2010; Pennestorfer and Schultz, 2012).

As gastrulation approaches, the exact direction of cell divisions are more difficult to track in all groups of lophophorates, and the symmetry of the biradial pattern is lost. Gastrulation occurs when the embryo has a very large number of cells (over a thousand in Brachiopoda and around hundred in Ectoprocta and Phoronidea) (Strom, 1977; Gruhl, 2010; Hejnol. 2010). Then, the resulting coeloblastula usually undergoes gastrulation by invagination, being asymmetric cell divisions commonplace during this process (Wood, 2007). In ectoprocta, however, epiboly is also observed along with invagination because the blastula is much more compact (stereoblastula) (Conklin, 1902; Wood, 2007). in lophophorates, the blastopore becomes the anus and the mouth forms secondarily as in deuterostomes, but this and other features (like the similarity between the lophophore and the tentacular crown of pterobranch hemichordates) seem to be a case of homoplasy (Passmaneck and Halanych, 2006; Laumer et al. 2015).

Cell fate maps in lophophorata are poorly known (Nielsen, 2002; Hejnol, 2010; Fuchs et al., 2011), but, since the variability in blastomere arrangement shown in lophophorata prevents the early determination of the D-quadrant via inductive cellular interactions (as in other spiralians, see previous section), the mode of cell fate determination seems to be indeterminate (Freeman and Lundelius 1992; Freeman, 1999; Hejnol, 2010; Pennestorfer and Schultz, 2012). That way, despite the clear symmetry planes of the biradial cleavage pattern, its symmetry axes are not always related to the adult body axes (Freeman, 1999; Nielsen, 2002). However, more detailed experiments (including cell-tracking) are required in this issue.

3.4.2.7. Entoprocta

Entoprocta (aka Kamptozoa) are small, sessile and often colonial aquatic animals, which have a characteristic filtering apparatus composed of a crown of ciliated tentacles (Wanniger, 2015). Entoprocta have affinities with both mollusks and annelids (trochozoa) and Bryozoans (Lophotrochozoa), but their belonging to one or the other superphylum is still debated (Nielsen, 2002; Hausdorf et al., 2007; Merkel, 2012; Nielsen, 2012; Wanniger, 2015). They superficially look like Bryozoa, but contrary to them, entoprocts have the anus inside the crown of tentacles, as long with the mouth (Rupert and Barnes, 1994; Nielsen, 2002). Other features, including the cleavage pattern, also differ.

Their cleavage, which has been only characterized in few genera (mainly *Loxosomella* (Merkel, 2012), *Pedicellina* (Hatschek, 1877; Dublin, 1905; Marcus, 1939) and *Barentsia*

(Malakhov, 1990)), is holoblastic, asynchronous, equal (cell divisions are symmetric) and more importantly, spiral (Rupert and Barnes, 1994; Merkel et al., 2012).

Eggs are yolky and spheroidal, and they present two polar bodies after fertilization: one in the animal pole and the other shifted about 90 degrees relative to the first one (Merkel, 2012). A third polar body appears in the 4-cell stage in the periphery of the blastula.

After the 4-cell stage the cell divisions lose the synchronization progressively, and the blastula becomes more flattened along the AV axis (Merkel et al., 2012). Despite that, in most of species (not in *Loxosomella* (Nielsen, 2012)), the blastula displays the typical features of the quartet spiral cleavage pattern. These features include the left/right alternation of each tier of four blastomeres, the size differences between micro- and macromeres (see section 3.4), and the pattern of cell-cell contacts between specific blastomeres (which resembles those found in mollusks and annelids, although it does not imply a phylogenetic relationship (Jenner, 2003; Merkel et al., 2012)).

During gastrulation, which starts at 100-cell stage approximately, the blastula (now coeloblastula) recovers its original spheroidal shape and then elongates along the AV axis, giving rise to a trochophore-like larvae mainly by means of cell rearrangements (Marcus, 1939; Nielsen, 1971; Merkel et al., 2012). Despite their cell fate map is difficult to characterize because all blastomeres are very similar, at least in *Pedicellina* it is similar to spiralian cell fate map (Marcus, 1939; Wanniger, 2015).

3.4.2.8. Cycliophora:

Cycliophora is a phylum of microscopic, sessile filtering animals, which were recently discovered in the mouthparts of the lobsters (Funch and Kristensen, 1995). Its phylogenetic affinities of are still discussed, but they seem to be the closest relatives of Ectoprocta and Entoprocta (Funch and Kristensen, 1995; Winnepennickx et al. 1998; Wanniger and Neves, 2015). The only genus of the phylum (*Symbion*), exhibits one of the most complex life cycles known in animals (Wanninger and Neves, 2015), involving asexual (budding) life stages and a trochophore-like larva.

When they reproduce sexually, a single embryo develops inside the female, occupying a relative large volume of her body (Neves et al., 2012). At least upon the 8-cell stage, cleavage seems to be holoblastic, with four macromeres and four micromeres. The blastomere arrangement bears no similarities to a spiral cleavage pattern; rather it resembles the polyaxial cleavage found is some poriferans (see section 3.1). Polar bodies have not been observed. Since this pattern has been

only observed in a settled specimen under transmission light microscopy, more data are required to characterize unambiguously the cycliophoran cleavage pattern.

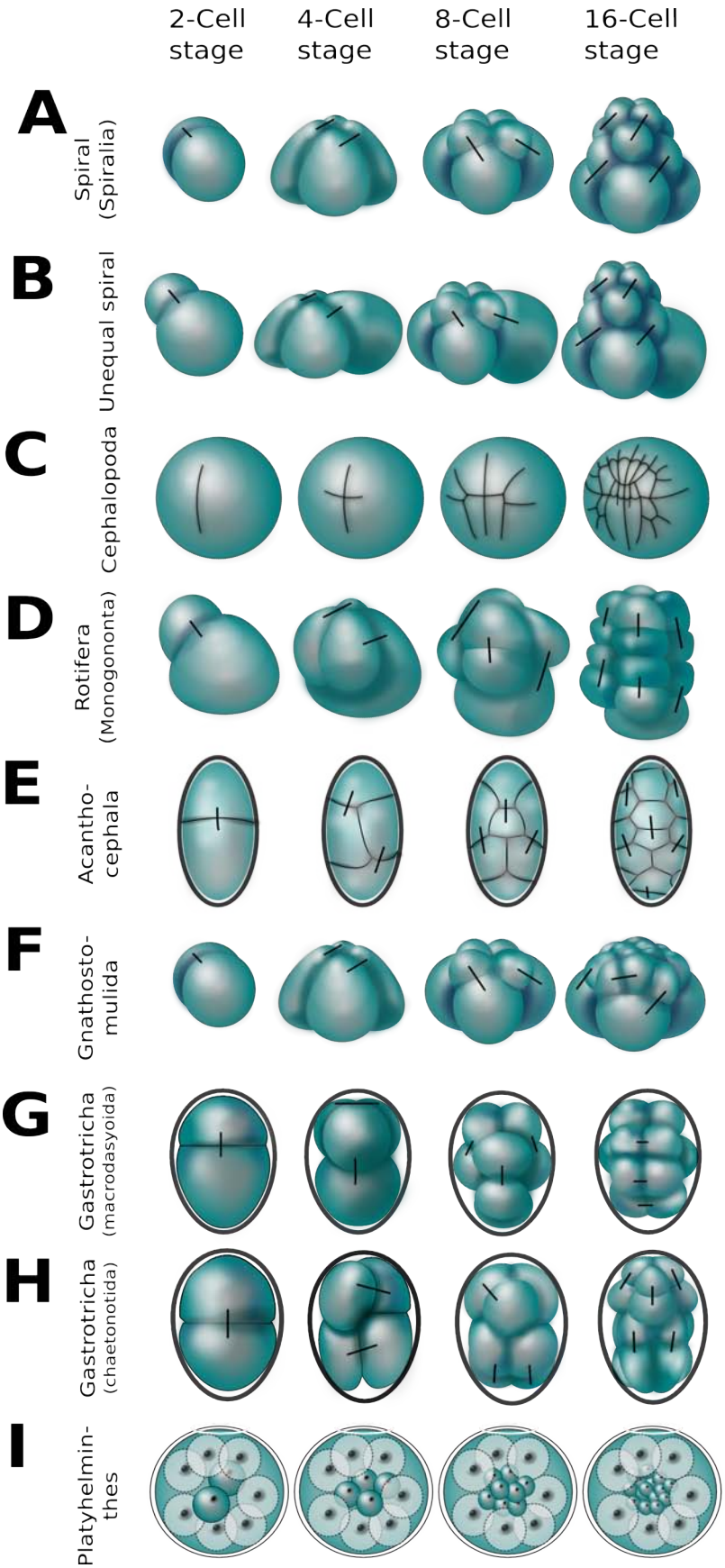


Figure 3. (From previous page) Example cleavage patterns found in the superphylum *Spiralia*. Most of them may be explained by means of the combination of a few conserved processes (see text). All blastulae are displayed in lateral view with the animal pole on the top, and the small straight lines link sister blastomeres when both of them are visible and identifiable. A) Equal "quartet" spiral cleavage pattern as displayed by many Spiralian taxa (the "pseudospiral" cleavage pattern found in other non-spiralian taxa is similar to this one until the 8-cell stage). B) Unequal "quartet" spiral cleavage pattern as displayed by many Spiralian taxa with determination of the D-blastomere (the biggest one, in the right side of the drawn blastulae) by asymmetrical segregation of cytoplasmic determinants. C) Meroblastic cleavage of Cephalopoda (mollusca). D) Idiosincratic cleavage pattern of monogononta (rotifera) (Eggshell not shown). E) Idiosincratic cleavage of Acantocephala. F) Spiral-like cleavage pattern of Gnathostomulida. G) Idiosincratic cleavage pattern of Gastrotricha (class Macrodasyoida). H) Idiosincratic cleavage pattern of Gastrotricha (class Chaetonotida). I) Disperse cleavage (*blastomerenanarchie*) of Platyheminthes

3.5. Chaetognatha.

Chaetognatha or arrow worms are small semitransparent rod-shaped planktonic predators (Brusca and Brusca, 1990). Its phylogenetic position within metazoa is among the most controversial ones: they have been placed within (Nielsen, 2012; Harzsch et al., 2015). Even its placement within either protostomes (ecdysozoa and spiralia/lophotrochozoa) or deuterostomes remains unclear, mainly because many of their genomic and anatomic features are not shared with any other animal taxa (autapomorphies), making comparisons very difficult. In here, we succinctly describe their early development without discussing their evolutionary affinities.

Eggs are rounded and, after fertilization either they are laid in the seawater in some species (genus *Saggitta*) or they remain some time inside the mother body (genus *Spadella*). In many cases, the oocyte has a dense cytoplasmatic granule (presumed to be a germ plasm determinant), which is sorted in one daughter cell in each cell division, remaining in the most vegetal cell of the blastula, at least until 64-cell stage (Carré et al., 2002).

In general, the cleavage is total (holoblastic), synchronous and equal (symmetric cell divisions), and it has been classically considered to be radial, as in deuterostomes (Brusca and Brusca, 1990). However, more recent studies focused on cell fates have shown that their development is more similar to protostomes (Hartzch et al., 2015).

The first two cell divisions are meridional, with the four resulting blastomeres (two more animal and two more vegetal cells) in a tetrahedral arrangement, each one in the vertex of an imaginary tetrahedron. The two more vegetal cells are connected by a great cross furrow, as so do the two more vegetal cells. Each cross furrow is perpendicular to the other when viewed from the animal pole (actually, this is a "rotational cleavage", thought this nomenclature is not used in

bibliography). The obliqueness of the pairs of sister cells respect to the animal-vegetal axis in this 4-cell stage has lead to some researchers to interpret this cleavage as having some similarities with the quartet spiral one (Shimotori and Goto, 2001), and even to apply to it the Conklin nomenclature developed for spiralians (in this case, the D blastomere is the one containing the germ plasm determinant) (Hartzch et al., 2015). However, since the blastomeres are very compressed (lacking the rounded shape) it is likely that this configuration arises from packing constraints imposed by the eggshell.

After this stage the arrangement fate of blastomeres becomes less clear and variable between species, thus making difficult to attribute it to radial or spiral patterns. For instance, in the genus *Spadella* the next cell divisions are asynchronous (passing through 6 or 14-cell stages) and leads to an irregular pattern, whereas in the genus *Sagitta* the next cell division is meridional, and each pair or sister blastomeres are almost totally aligned with the animal-vegetal axis, as in radial cleavers (Elpatewsky, 1909). However, some spiral features like the left-right alternation of cell divisions respect to the AV axis that have been purported in *Sagitta* (Elpatewsky, 1909), and the cell fates suggest that the Chaetognatha early cleavage is more similar to protostomes than to deuterostomes (Wanninger, 2015). In both genera, by the 32-cell stage the blastomeres in the animal half of the embryo follow a radial-like pattern, whereas the vegetal-most blastomeres form sort of rosette around the vegetal pole (Hartzch et al., 2015).

In the gastrulation of chaetognatha, after the formation of the endoderm by an invagination process, a second opening forms the stomodeum in the opposite site as in deuterostomes. After that, both openings close and then are re-established some time after hatching. Thus, the blastopore does not become the future anus, as it does in deuterostomes (Hartzch et al., 2015). Summarizing, Chaetognatha possess a combination of a protostome-like early cleavage with a transitory deuterostome-like gastrulation process (Hartzch et al., 2015).

3.6. Deuterostomes.

Deuterostomes are an extremely diverse group at both morphological and taxonomical level (Brusca and Brusca, 1990). They ere distinguished from the rest of metazoan clades (protostomes) by their embryonic development; their first opening (blastopore) in their blastulae becomes the anus, while in protostomes, it becomes the mouth. Many works include within deuterostomes the lophophorates (Rupert and Barnes, 1994), or even the Xenacoelomorpha (Perseke et al., 2007), but nowadays it seems that Deuterostomes only include Echinodermata (sea urchins, sea stars and their relatives), Hemichordata (acorn worms) and Chordata (ascidians, amphioxius and vertebrates)

(Nielsen, 2012; Rouse et al., 2016;). In general, deuterostome cleavage is holoblastic, with loosely attached blastomeres and typically radial. In radial cleavage, early cell divisions follow Sachs´ rule: they are either parallel or perpendicular to the animal-vegetal axis, depending on their relative position along the animal-vegetal axis. Thus, along this axis blastomeres are always located one on the top of each other, not in oblique positions as in spiralia. In addition, some of these cell divisions are often asymmetric, yielding groups of cells of different size sorted along the animal-vegetal axis. Deuterostome taxa exhibit slightly different radial patterns, which arise from changes in the cell adhesion and in the timing and location of asymmetric cell divisions.

3.6.1.Echinodermata:

The echinoderms (starfishes, sea urchins and relatives) are a large phylum of exclusive marine and usually benthic animals which posses an internal skeleton made of calcareous ossicles and a multi-function network of fluid-filled canals derived from the coelom, called water-vascular system (WVS). The phylum Echinodermata has currently five recognized classes with penta-radial symmetry (crinoidea or sea lilies, ophuroidea or bristle stars, asteroidea or starfishes, echinoidea or sea urchins and holoturoidea or sea cucumbers) plus an even large number of extinct classes including penta-radial, tri-radial, bilateral and even asymmetric forms.

Although some species brood their embryos, most representatives of each class have a dispersive larva that is planktorophic in forms with small eggs and nonfeeding (lecitotrophic) in forms with large eggs fully provisioned with yolk (Raff and Byrne 2006) (See below for a brief discussion on echonoderm larval types).

Most developmental studies in echinoderms have been carried out in echinoids (specifically in the model species *Strongylocentrotus purpuratus*), for which the circuitry of endomesoderm specification has been exhaustively characterized. Its development is highly regulative, a feature that is extensive to all echinoderms (Rupert-Barnes, 1994; Gilbert and Raunio, 1997). Thus, we are going to describe the early development of the sea urchin embryo, explaining after the differences with other echinoderm groups.

In sea urchins, the cleavage is holoblastic, radial, and symmetric in all but the fourth cell division (Fig. 4C). In this fourth cell division round (8 to 16-cell stage), the four upper (animal) blastomeres divide meridionally forming equal-sized cells called mesomeres, but the lower four cells divide asymmetrically and equatorially, producing four larger macromeres and below them (in the vegetal pole) four smaller cells called micromeres (Summers et al., 1993). At the fifth cell division the eight mesomeres divide symmetrically and equatorially, forming two tiers of

blastomeres in the animal hemisphere, one staggered above the other. The four macromeres divide meridionally, forming a tier of eight cells, while the micromeres divide again asymmetrically, generating four large micromeres and four small micromeres (Okazaki, 1975; Cameron and Davidson, 1991). Then, at sixth cell division all blastomeres divide equatorially, producing the 60-cell stage coeloblastula. At the 128-cell stage the gastrulation properly starts with a hollow sphere of nearly epithelial cells (a blastoderm with cell junctions and extracellular matrix on both the interior and exterior surfaces). The central cavity (blastocoel) is formed by the adhesion of the blastomeres to the hyaline layer, and expanded by an active influx of water (Dan, 1960; Gilbert and Raunio, 1997).

As development proceeds, the animal pole of the blastula develops a tuft of cilia and the archenteron extends dramatically, sometimes triplicating the length of the embryo along the dorsoventral axis. In addition, several pairs of skeletonized arms appear and extend outwards the blastula surface, reaching that way the larval (pluteus) stage (Arnone et al., 2015).

During cleavage, in the process of cell specification, that is regulative, up to five territories (groups of cells with different gene expression patterns, cell lineage histories and cell fates) are distinguished (Davidson et al. 1998; Salazar-Ciudad, 2003): the small micromere, the skeletogenic mesenchyme, the vegetal plate, the aboral ectoderm, and the oral ectoderm territories (Davidson et al., 1998).

In asteroidea, crinoidea and holoturoidea, the cleavage pattern, cell fates and regulatory genes expressed are fairly similar to that of sea urchins, except for a few facts: all cell divisions (including the fourth one) are symmetric, the perpendicularity between consecutive cell divisions (and thus the radial pattern) are not so clear as in sea urchins, many of them (crinoidea and asteroidea) have large transient pore in the vegetal pole of th blastula, and they have no skeletogenic elements in the pre-larval stage as sea urchins do, because the larvae of starfishes (bipinnaria) and holoturoidea (auricularia) do not possesses skeletonized arms (Holland, 1991).
Furthermore, in some species of sea cucumbers (*Parastichopus parvimensis*) cell divisions are not synchronous and blastomeres show little cell-cell adhesion, causing gaps between them.

The cell divisions in ophiuroidea also are symmetric and synchronous, but, at least in some species (*Ophiopholis aculeata*) the cell fate map indicates that ecto- and endomesoderm are already established at very early stages, posting a major difference with other echinoderms, in which they are established lately via inductive interactions (regulative development). In addition, in this bristle-star the cleaving embryo shows a very compact appearance (likely caused by an increased cell-cell adhesion) in which the tiers of blastomeres are almost no recognizable (Arnone et al., 2015).

All classes of echinoderms have examples of both planktotrophic and lecitotrophic larvae

depending on the species (except Crinoidea which has only lecitotrophic non-feeding larvae). Whereas the non-feeding larvae have all a barrel-like appearance irrespective of the class they belong to, each class has its own typical morphology of feeding larvae. Echinoidea and ophiuroidea have a larvae with long stiff appendages (called respectively echinopluteus and ophiopluteus); whereas asteroidea and holoturoidea have larvae with a contorted ciliary grooves (called respectively bipinnaria and auricularia). These latter are very similar to the tornaria larva of the Hemichordata (see next section) and are considered to represent the basal-type larva for both phyla (which form the superphylum Ambulacraria) (Rupert and Barnes, 1994; Raff and Byrne, 2006).

The visual resemblance of the types of larvae between echinoderm classes does not fit the phylogeny of the group, and several theories have been proposed to explain this seeming discrepancy between adults and larvae (Sly et al. 2003; Williamson, 2009; Minelli, 2010). However, the arms and ciliary structures in the different types of larvae are not necessarily homologous between them, and it seems that larval phases have arisen through intercalation between the gastrula and juvenile life phases. Under this scenario, the resemblance between the larvae of distantly related echinoderm species represents a case of convergent evolution due to common life strategies and common underlying developmental mechanisms (Wray, 2000; Sly et al. 2003; Salazar-Ciudad, 2010). In addition, lecitotrophic larvae (or *vitellaria*, which have a much simple pattern of ciliation) appear to have arisen independently and frequently in many echinoderm clades from plaktotrophic larvae (Arnone et al., 2015).

After the free-swimming planktonic existence, the bilateral larvae undergo a metamorphosis to a radially-symmetric adult organisms. This change is of great interest and there are many reviews on echinoderm metamorphosis and the morphological changes associated with these later developmental events, but a detailed description of them is beyond the scope of this review (Burke 1989; Arnone et al., 2015).

3.6.2. Hemichordata:

Hemichordates comprise the class Enteropneusta (aka acorn worms) and the class Pterobranchia (small semi-sessile worms equipped with a lophophore-like feeding apparatus).

Most enteropneust have small nonyolky eggs that develop into a plaktotrophic ciliated larvae (tornaria), but a few groups of direct developers, such as species of *Saccoglossus* have large yolky eggs that develops into a lecitotrophic (and very short-lived) larvae (Kaul-Strehlow and Rottinger, 2015). In enteropneusta, differences between direct and indirect developers only concerns later developmental stages and the relative speed of development, being the cleavage patterns and

fate maps identical. As in sea urchins, blastomeres in the eight-cell stage are arranged in a cubic-shaped configuration, and by the 16-cell stage the blastula exhibit an animal tier of eight cells (anterior ectoderm), an “equatorial” tier of four cells (posterior ectoderm) and, directly below, a vegetal tier of four cells (endomesoderm) (Tagawa et al. 1998). Subsequent cell divisions give rise to a a rounded coeloblastula, which in the direct developers is deformed along the A-V axis in order to acquire an elongated shape better suited to swimming (Colwin and Colwin, 1953).

In pterobranchia, the cleavage pattern is also radial but it is somehow disturbed because the 4-cell stage acquires a diamond-shaped rather than the square-shaped configuration (most likely due to an increased cell-cell adhesion). In turn, the next cell division is very asymmetric, so that the animal micromeres appear in two groups of two (some authors call this configuration bilateral cleavage (Rupert-Barnes, 1994)). The cleavage pattern in the next cell-stages is not clear but gastrulation is known to be by ingression or delamination, after which a planktotrophic ciliated larvae arises (Lester, 1988).

3.6.3.Chordata

The phylum chordata includes three subphyla: urochordata, cephalochordata and craniata (incl. Vertebrata), although some authors suggest that each one must be a phylum on its own (Nielsen, 2012; Sathoh et al., 2014). Their distinctive feature is the presence of a notochord, a hollow dorsal nerve cord, pharyngeal slits and, at least in some stage of their life cycle, post-anal tail (Rupert-Barnes, 1994).

In ascidians (Urochordata), before fertilization the egg is polarized along the animal-vegetal (AV) axis: endoplasmic reticulum-rich cortex (cER), associated maternal mRNAs and mitochondria are almost absent in the animal pole (Prodo et al., 2008). The radial holoblastic cleavage is replaced by another pattern called bilateral cleavage, that is remarkably conserved, even between distantly related species (eg. *Ciona intestinalis* and *Halocynthia roretzi*). In it, the first cell division is symmetric and occurs along the AP axis, partitioning the embryo into bilaterally symmetric left and right blastomeres (the furrow between these two blastomeres establishes a plane of symmetry which will continue separating the future right and left halves of the embryo). The second cleavage occurs perpendicular to the first, and the third cleavage occurs perpendicular to both the first and the second (Sachs' rule), resulting in a cubic-shaped eight-cell stage. Each of the four blastomeres in the right and left halves are called a “line” or a “quadrant” as in spiralians (but in this case the quadrants are not arranged perpendicular to the AV axis but parallel to it). From the most animal part to the most vegetal part of the embryo, the quadrants are named “a, A, b and B”.

Since the cleavage pattern is constant a special Conklin nomenclature is used to label unequivocally individual blastomeres (Conklin, 1905).

After the 8-cell stage, Sachs´ rule makes way for other forms of oriented cell division, and each quadrant gives rise to a multitude blastomeres organized according to a unique and invariant cleavage bilateral pattern (the right half of the embryo is kept as the mirror image of the left one). As cleavage goes on, some cell divisions are not longer symmetric. These asymmetric cell divisions (mediated by an actin-rich structure called the centrosome-attracting body (CAB) which attracts the centrosome to the posterior cortex) enable the asymmetric segregation of maternal determinants to one of the two daughter cells (Munro et al. 2006).

In turn, these asymmetric segregation of maternal determinants, combined with inductive signals between neighboring cells are involved the cell fate determination since early developmental stages. In this interplay between inductive and cell-autonomous mechanisms, the timing, orientation and asymmetry cell divisions are relevant factors because they determine altogether the geometry and extension of cell-contacts among inducing cells (below a certain threshold area, cells are unable to respond). Conversely, inductive signals can shape the timing and orientation of cell divisions, thus paving the way for subsequent fate decisions (Darras and Nishida, 2001; Munro, 2006). This early cell-fate determination has been argued to represent an adaptation to development with a small number of cells (Munro, 2006).

Because of this interplay, the ascidian cleavage is a very suitable system to study the dynamic coupling among gene networks, cell fate allocation and cell geometry (some lines of research are exploiting this model by means of computational modeling and advanced 3D imaging) (Tassy et al., 2010).

Early gastrulation starts when the coeloblastula arrives at the 110-cell stage, the cell fate map (and the germ layers) have been established. Gastrulation is driven by invagination, which happens in a two-step process: an initial apical constriction followed by a basolateral cell shortening (Sherrard et al. 2010). Then, the posterior part of the blastula suffers a great elongation by cell intercalation, thus generating the tail rudiment of the tadpole larva (Munro and Odell, 2002).

In cephalochordata (*Amphioxus* or lancelet), the cleavage was long believed to be very similar to Urochordata, but recent studies show that the cleavage pattern is not bilateral but fully radial, and close to Echinoderms' one (Cerfontaine, 1906; Holland and Yu, 2004). In addition, since the XIX century, the mode of cell fate determination was known to be regulative and not determinative (mosaic) as in ascidians (Wilson, 1893). The origins of this misinterpretation is twofold: the difficulty of classical microscopists to cope with the tiny size of amphioxus eggs

(<140μm) and the the ideological prejudice arising from of the close relationship between amphioxus and tunicates (Holland, 2015). Main developmental studies come from the species *Branchiostoma floridae, B. belcheri* and *B. lanceolatu,* in which the cleavage process is a follows:
Just after fertilization, the sperm pronucleus migrates to the vegetal pole, and, as typical radial-cleavers, the the first cell divisions proceed at right angles one to another and respect to the AV axis. Cleavage is holoblastic, and blastomeres are roughly equal in size (vegetal blastomeres slightly larger) and very rounded (not tightly adherent) until the mid-late blastula.

One (or two in *B. belcheri*) of the vegetal-most blastomeres becomes the germ-line precursor, a role that is achieved by the segregation (during cleavage) of the “pole plasm”: a part of the egg cortex which presents a high density of endoplasmic reticulum and associated germ cell marker RNA (maternal *Vasa* and *Nanos)* (Wu et al. 2011).

Concerning gastrulation process, classical works also show a disagreement about the formation of the blastula (Holland, 2015). Fortunately, subsequent work has shown that gastrulation is far simpler than most of these earlier author believed : blastula simply gastrulates by invagination, with slight movement of cells over the gastrula's lips. In the emerging gastrula, the cell layer covering the archenteron is much wider (columnar epithelium) than the outer layer, and the blastopore constitute the posterior end of the embryo (Zhang et al. 1997).

Other important departures from radial cleavage are found within vertebrates, and many of them are driven by a great amount of yolk in the vegetal pole (Gilbert and Raunio, 1997). In amphibians, this causes the equatorial cell divisions to be displaced towards the yolk-free animal pole (displaced radial cleavage). If the yolk is distributed over all the egg, as in fishes, reptiles and birds, only a meroblastic cleavage restricted to the surface of the animal pole can happen. In this case, first cell divisions still follow Sachs' rule as in radial cleavage, but the stereotypic pattern disappears soon. This kind of cleavage is called discoidal, and presents morphological commonalities with the one found in some Arthropoda.

In placental mammals, early development takes place the mother´s body in a close metabolic dependence. Their eggs are consequently yolk-free, and the cleavage is holoblastic. The second round of cell division is not only perpendicular to the previous one, but also perpendicular between the two blastomeres (one is meridional, the other equatorial). Because of the resulting tetrahedral configuration, it is called rotational cleavage (notice the resemblance with the rotational cleavage of nematoda is restricted to the 4-cell stage, and the mechanisms involved differ). After that, cell divisions become asynchronous with cycles of increased cell adhesion, and their directions are

determined by Hertwig´s rule and the contacts between adjacent blastomeres, losing any radial (or even regular) appearance.

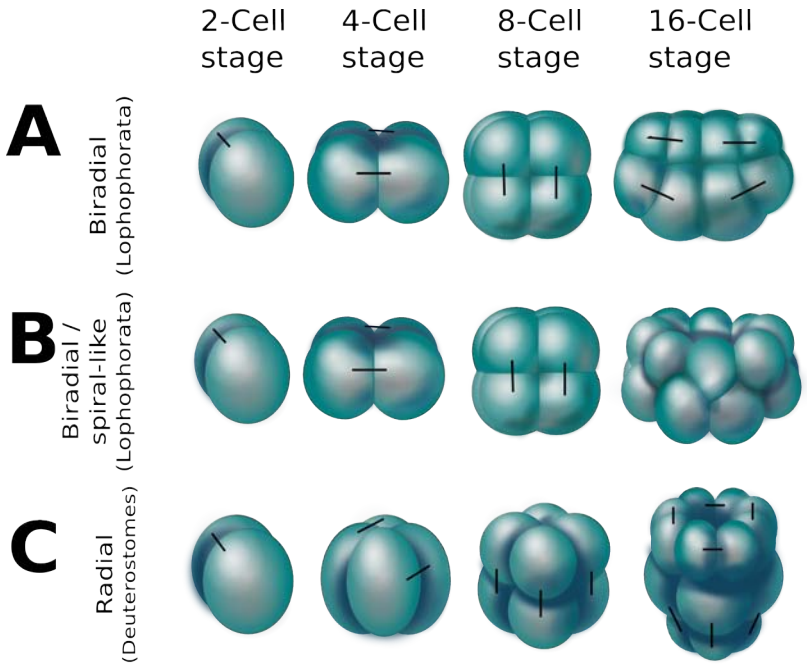


Figure 4. Radial and radial-like cleavage patterns. Most of them may be explained by means of the combination of a few conserved processes (see text). All blastulae are displayed in lateral view with the animal pole on the top, and the small straight lines link sister blastomeres when both of them are visible and identifiable. A) Biradial cleavage pattern of some Lophophorates. B) A lophophorate biradial pattern with transient spiral-like features in the 16-cell stage. C) Radial cleavage pattern of a Deuterostome (sea urchin).

4. Conclusions.

Despite their diversity, most cleavage patterns seem to have been built by means of the combination of a handful of similar and evolutionary old cell processes (Salazar-Ciudad et al. 2003). Among the many developmentally available patterns, most metazoans seem to use those exhibiting both mechanical stability and basic symmetry axes. These invariant cleavage patterns, in which the timing, orientation and symmetry of cell divisions are precisely defined for all cells, allow the very early use of inductive mechanisms for cell fate and body axes determination. This early cell-fate determination is thought to be adaptive for those metazoans which develop fast in a free-swimming, planktotrophic larvae (e.g. most Spiralians). These taxa spawn large amounts of small and relatively yolk-free eggs in which the blastomeres are induced to form larval organs (which are made of very few cells) at very early developmental stages (Salazar-Ciudad, 2010). Thus, these early inductive events take place in the context of a few cells (usually through precise short-ranged signalling), which requires a constant blastulae architecture (Wray, 2000; Salazar-Ciudad, 2010). Otherwise, the variable arrangement of blastomeres (or even the mispalcement of a single cell) would lead to a abnormal inductive events, and consequently to catastrophic changes in the larval morphology. In some cases (e.g. ascidians), the precise signalling between cells that takes place within a constant blastulae architecture is not only used for cell fate determination, but also for the specification of the timing and orientation of the next cell divisions, in a dynamic interplay between morohogenetic and inductive mechanisms.

In other groups (e.g. vertebrates, cnidarians), whose relative blastomere neighbourhood is very variable, so that inductive events associated with organogenesis can only take place much later in development, involving large number of cells, long-ranged signalling between them, and in many cases extensive cell movements (Salazar-Ciudad, 2010). Many of these groups with a loose control of the cell processes during early cleavage exhibit similarities in the blastomere arrangement in the very early stages, that seem to be due to mere packing principles.

Evolutionary transitions between different cleavage patterns have happened many times. Some of them may be explained by adaptive changes in their underlying cell processes (e.g. changes in cell adhesion, in the amount of yolk or the acquisition of a more rigid eggshell), but the developmental bases of these transitions are poorly understood (Wray, 2000). Much work remains to be done (involving experiments in model and non-model species, and computational approaches) in order to disentangle how the interplay between developmental and selective forces have sculpted the geometry of metazoan cleavage patterns.

TAXON	TYPE	YOLK	DETER-MINATION	VARIATION	DIRECTION				SYMMETRY				SYNCHRONY				GASTRULATION PROCESS	SHELL	MAIN FEATURES
					1	2	3	4	1	2	3	4	1	2	3	4			
Porifera (Halisarcida, Demospongiae)	Polyaxial	HOL	N?	Y	M	M	GE	GO	Y	Y	Y	Y	Y	Y	Y	Y	Difficult to identify (Blastula-like larva)	Brood Chamber (G)	Cell divisions perpendicular to cell-surface. Hollow blastula of cone-shaped blastomeres.
Porifera (Hexactinellida)	Radial-like	HOL	N?	?	M	M	E	GO	Y	Y	Y	Y	Y	Y	Y	Y	Difficult to identify (Blastula-like larva)	Brood Chamber (G)	Hexactinellida and demospongiae. Divisions after 8-cs perpendicular to embryo Surface, hollow blastula.
Porifera (Oopscas ,Hexactinellida)	Pseudospiral	HOL	N?	Y	M	MO	EO	RO	Y	Y	Y	N	Y	Y	Y	Y	Difficult to identify (Blastula-like larva)	Brood Chamber (G)	Yolk-rich macromeres in the centre, surrounded by micromeres (syncityum) . Transitory stage of chaotic cleavage?
Porifera (Homoscleromorpha)	Chaotic / Anarchic	HOL	N	Y	R	R	R	R	Y	Y	Y	GY	N	N	N	N	Difficult to identify (Blastula-like larva)	No	Homoscleromorpha and ovoviviparous demospongiae.
Porifera (Calcaronea, Calcarea)	Incurvational	HOL	N?	GN	M	M	M	M	Y	Y	Y	Y	Y	Y	Y	Y	Difficult to identify (Blastula-like larva)	Incubation Chamber	Epithelioid. Tetraradially-symmetric when viewed from the animal pole.
Cnidaria (Hydrozoa and Anthozoa)	Chaotic / Anarchic / Pseudospiral	HOL	N	Y	R	P	R	R	YG	YG	YG	YG	YG	YG	YG	YG	Coeloblastula. Invagination.	No (G)	Great intraspecific variation. Transitory pseudospiral. The more important the Sedentary stages in postembryonic development, the more anarchic the cleavage.
Cnidaria (Scyphozoans)	Chaotic / Anarchic / Radial-like	HOL	N	Y	R	P	P	R	Y	Y	Y	Y	Y	Y	Y	Y	Coeloblastula. Invagination in the animal Pole. Planula larva.	No (G)	Transitory fusions and syncitiums. First Furrow correlates with the invagination site of larvae.
Placozoa	Pseudospiral	HOL (fat droplets)	N	N	?	?	?	?	Y	Y	Y	Y	Y	Y	Y	Y	?	Y (Fertilization Membrane)	Tetrahedral in 4 cell-stage. No clear directions of cell divisions. No beyond 64-cs in culture
Ctenophora	idiosynccratic (Bilateral-like)	HOL	Y	N	M	M	OE	M	Y	Y	N	N	Y	Y	Y	Y	Epiboly	No	Tetra-radial symmetry from 4-cell stage. Blastula symmetry axes correlate with adult body axes.
Acoela	Duet spiral	HOL	N	N	M	O	O	O	Y	N	N	N	Y	Y	Y	Y	16 cell-stage. Internalization of Endomesodermal Precursors 3A and 3B.	No	Differences with spiral pattern seem to arise from the relative timing of cell divisions and cell movements. Micromeres CCW.
Nemertodermatida	Latitudinal	HOL	Y	N	M	E	M	M	Y	N	N	N	Y	Y	N	N	64 cell stage. Internalization.	Y	Similar to duet spiral. Micromeres CW.
Priapulida	Radial-like	HOL	?	?	M	M	E	E	Y	Y	YG	YG	Y	Y	Y	Y	64 cell stage. Internalization by Cell migration.	Y?	Very regular and symmetric. Stereotypic. Sachs' like rule on blastomeres' surface.
Kynorhyncha	Radial-like	HOL	?	?	M	?	?	?	Y	Y	Y	Y	G	G	G	G	16 cell stage. Unknown process.	Y?	Blastomeres organized into "layers". Similar to priapulids.
Nematoda (Non-enoplids)	Rotational	HOL	Y	Y (IS)	M	MO	EO	V	N	N	N	N	Y	Y	Y	N	24 cell stage. Internalization	Y (Spheroidal)	Not very variable. Direction of cell divisions driven by several mechanisms acting together (see text for detaisl)
Nematoda (Enoplia)	Pseudospiral	HOL	Y	Y (IS)	M	V	V	V	NG	NG	NG	NG	YG	YG	YG	YG	64 cell stage	Y (Usually very Elongated)	Blastomere arrangement correlates with eggshell shape.
Nematomorpha	Pseudospiral	HOL	N	Y	M	M	E	E?	Y	Y	Y	Y	G	G	G	G	?	?	After 2-cell stage very variable. After 16-cell stage, Pseudospiral pattern vanishes.
Arthropoda (Hexapoda)	Intralecithal / Superficial / Centrolecithal	MER (HOL in some Collembola)	N	Y	R	R	R	R	Y	Y	Y	Y	YG	YG	YG	YG	Short, intermediate or Long germband formation.	Y (Usually very Elongated)	Energids migrate from the center of the embryo to the peryphery to form the perypheral blastoderm (later in holometabolous)
Arthropoda (Chelicerata)	Discoidal / Chaotic	MER (HOL in mites And scorpions)	N	Y	R	P	P	P	Y	Y	Y	Y	YG	YG	YG	YG	Short to intermediate Germ band formation.	Y (Usually very Elongated)	Square configurations until the 16-cell stage. Sachs-like rule in the embryo surface.
Arthropoda (Crustacea)	Spiral-like	HOL	Y	N	M	M	EO	EO	Y	Y	Y	Y	YG	YG	YG	YG	Short to intermediate Germ band formation.	Y (Usually very Elongated)	Cell fate map, mode of cell-fate determiantion and cell processes involved (e.g. cell-contact rule) are not the same as in Spiralia.
Arthropoda (Myriapoda)	Intralecithal / Discoidal	HOL (G) / (MER in Chilopoda)	?	?	M	M	?	?	Y	Y	N?	N?	Y	Y	N	N	Short to intermediate Germ band formation.	Y (Usually very Elongated)	In some cases energids migrate radially to the peryphery (intrapyramidal cleavage, see text)
Onychophora	Discoidal (Viviparous) / Irregular (Oviparous)	MER (Viviparous) / HOL (Oviparous)	?	?	R	P	P	P	Y	Y	Y	Y	Y	Y	Y	N	Invagination / Germ band formation	YG	Spiral-lie (annelid-like) cell-fate map.

TAXON	TYPE	YOLK	DETER-MINATION	VARIATION	DIRECTION				SYMMETRY				SYNCHRONY				GASTRULATION PROCESS	SHELL	MAIN FEATURES
					1	2	3	4	1	2	3	4	1	2	3	4			
Tardigrada	Pseudospiral	HOL (Yolk droplets)	?	Y	M	M	P	P	Y	Y	N	?	Y	Y	Y	?	?	Y Sub-spherical)	Very hard eggshells. Contradictory data concerning the cleavage pattern.
Spiralia (equal cleavers)	Equal (quartet) spiral	HOL	Y / N	N	M	M	O	O	Y	Y	GN	GN	Y	Y	Y	Y	Epiboly	V	Late determination of the 4D blastomere through cell interactions (inductive mechanisms, regulative)
Spiralia (unequal cleavers)	Unequal (quartet) spiral	HOL	Y	N	M	M	EO	EO	N	N	GN	GN	Y	Y	Y	Y	Epibloy	V	Early determination of the 4D blastomerethrough asymmetric seggregation of Maternal determinants (autonomous mechanism, determinative).
Cephalopoda	idiosynccratic	MER	Y	N	M	M	EO	EO	Y	Y	GN	GN	Y	Y	Y	Y	?	V	Total lack of spiral features
Rotifera (Monogononta)	Spiral-like	HOL (around nucleus)	Y	N	M	M	E	E	N	N	N	N	Y	Y	Y	Y	Epiboly	Y (Multi-layered)	A spiral-like one, very asymmetric from the beginning. No alternation. Tower-like shape at the end.
Acanthocephala	Distorted spiral	HOL	?	?	E	EO	EO	EO	N	N	N	N	Y	Y	Y	N	?	Y (Very Elongated)	Very elongated eggshell causes meridional divisions to result in oblique arrangements. Bilaterally symmetric cleavage pattern
Ganthostomulida	Spiral-like	HOL	?	?	M	M	E	E	Y	Y	Y	N	N	Y	Y	Y	?		Spindle alternation. Micromeres CW. Annelid-like cross. Only one species decribed *(G. jenneri)*.
Gastrotricha (Chaetonotida)	Idiosyncratic. Some spiralian features	HOL	Y	N	E	M	M	E	Y	Y	Y	YG	Y	Y	Y	Y	32 cell stage. Interiorization of lateral A-blastomere descendants	Y (Elongated)	Blastomere rotation in the 4 cell-stage.
Gastrotricha (Macrodasyoida)	Idiosyncratic. Some spiralian features	HOL	Y	N	E	MV	E	M	N	YG	YG	YG	Y	Y	Y	Y	30 cell stage. Interiorization of lateral D-blastomere descendants	Y (Elongated)	Rearrangement of the 3 anterior blastomeres in the 4-cell stage cells. Bilterally symmetrical.
Platyhelminthes	Idiosynccratic (perypheral Yolk cells involved)	HOL (free-living) / MER (parasitic)	N	Y	GR	GR	GR	GR	Y	YG	YG	YG	Y	YG	YG	YG	Epiboly (G) / idiosynccratic	Egg capsule / Hull membrane	Yolk cells (vitellocytes) usually are perypheral at the beginning, belng internalized Later. Hull membrane is a blastomere derivative, not maternally provided.
Entoprocta	Spiral-like	Yolky (but still HOL)	Y	N	M	M	EO	?	Y	Y	Y	N	Y	Y	N	N	100-cell stage. Invagination.	N ?	Cell-fate difficult to characterize, but similar to Spiralia. CW. Lenticular blastula. Very low cell adhesion.
Ectoprocta	Biradial	Yolky (but still HOL)	N?	Y	M	M	E	M	Y	YG	YG	YG	Y	Y	Y	Y	64-cell stage. Invagination / Epiboly	Gonozoid	Polyembriony. 16-cell stage four tiers of four blastomeres in line. Symmetry axes not related to those of the larva. Very low cell adhesion.
Brachiopoda	Radial-like	HOL	N?	Y	M	M	E	EO	Y	Y	Y	Y	Y	YG	YG	YG	1000-cell stage. Invagination.	N?	Radial-like pattern quickly vanishes, elongated blastomeres with central blastocoel
Phoronidea	Spiral / Radial / Bilateral	HOL	N?	Y	M	M	E	M	Y	Y	Y	Y	Y	Y	Y	Y	Invagination	N?	Conflicting data regarding cleavage pattern. In 16-cell stage four paralell rows of f our cells each one. Very Variable. Alternating (CW-CCW) cell divisions as in spirals.
Cycliophora	Spira l Polyaxial-like	HOL	?	?	M	M	?	?	Y	Y	N	N	Y	Y	Y	?	?	Y (Mother Body)	Only one specimen. Not well characterized.
Chaetognatha	Rotational-like	HOL	?	Y (IS)	M	M	PO	PO	Y	Y	Y	Y	Y	Y	YV	YV	32-cell stage. Inmigration	Y (Rounded)	Tightly packed blastomeres. Protostome-like early cleavage and Deuterostome-like late cleavage.
Echinodermata	Radial	HOL	N	N	M	M	E	ME	Y	Y	Y	N	Y	Y	Y	Y	128-cell stage. Coeloblastula. Ingression.	Y (Hyaline Layer)	Regulative cell fate determination. Micro-, meso- and macromeres in the 32-cell stage.
Hemichordata	Radial	HOL	?	?	M	M	EG	EG	Y	Y	Y	NG	Y	Y	Y	Y	Coeloblastula. Ingression.	?	Radial pattern disturbed in Pterobranchia because the diamond-shaped configuration (not coplanar) in the 4-cell stage.
Chordata (Urochordata)	Bilateral	HOL	Y / N	N	M	P	P	V	Y	Y	Y	N	Y	Y	Y	YG	110-cell stage. Coeloblastula. Two-step invagination (see text)	N	Sachs-like rule in first cell divisions. Directions of cell divisions are determined by Both inductive and cell autonomous mechanisms.
Chordata (Cephalochordata)	Radial	HOL	N	N	M	M	E?	?	Y	Y	Y?	Y?	Y	Y	Y	Y	Invagination	N	Regulative cell fate determination. Contradictory interpretations of the cleavage pattern.
Chordata (Vertebrata)	Rotational / Discoidal / Idiosynccratic	HOL	N	Y	M	GO	V	V	Y	Y	Y	Y	Y	Y	Y	Y	V	V	Typical radial pattern is usually lost by adaptations arising from viviparism (e.g. placental mammals) or very yolky eggs (e.g. birds)

Table 1 (previous pages). Summary of the main features of the cleavages of the main metazoan taxa. Abbreviations: ?- Not conclusive or conflictive data. CCW- Counter-clockwise. CW- Clockwise. E- Equatorial. G- In general. HOL- Holoblastic. IS- Inter-specific. M- Meridional. MER- Meroblastic. N- No. O- Oblique. R- Random. V- Variable. Y- Yes.

BIBLIOGRAPHY:

Adamska, M., Degnan, S.M., Green, K.M., Adamski, M., Craigie, A., Larroux, C. and Degnan, B.M. 2007. Wnt and TGF-beta expression in the sponge Amohimedon queenslandica and the origin of metazoan embryonic patterning. *PloS One* 2(1):e1031.

Adamska, M., Degnan, B.M., Green, K. and Zwafink, C. 2011. What sponges can tell us about the evolution of developmental processes. *Zoology*. 114(1):1-10.

Aguinaldo, A.M.A., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R.A. and Lake. J.A. 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature*, *387*(6632):489-493.

Akiyama, M., Tero, A. and Kobayashi, R. 2010. A mathematical model of cleavage. *J. Theor. Biol.* 264:84-94.

Alberch, P. 1982. Developmental constraints in evolutionary processes. In: Bonner JT, ed. *Evolution and development*. Dahlem Konferenzen. Springer-Verlag: Heidelberg. P313–32.

Alwes, F. and Scholtz, G. 2004. Cleavage and gastrulation of the euphausiacean Meganyctiphanes norvegica (Crustacea, Malacostraca). *Zoomorph.* 123:125–137

Anakina, R.P. 1997. The cleavage specificity in embryos of the Barents Sea sponge *Leucoselenia complicata* Montagu (Calcispongiae, Calcaronea). In: Ereskovski AV, Keupp H, Kohring R., eds. *Modern problems of poriferan biology*. Berliner Geowissenschaftliche Abhandlung. Freie Universitat, Berlin 20:45-53.

Anderson, D.T. and Manton, S.M. 1972. Studies on the Onychophora. VIII. The relationship between the embryos and the oviduct in the viviparous placental onychophorans *Epiperipatus trinidadensis* (Bouvier) and *Macroperipatus torquatus* (Kennel) from Trinidad. *Philos. Trans. R. Soc. B Biol. Sci.* 264:161–189

Arnone, M.I., Byrne, M. and Martinez, P. 2015. Echinodermata. In *Evolutionary Developmental Biology of Invertebrates 6* (pp. 1-58). Springer Vienna.

Aw, S. and Levin, M. 2009. Is left-right asymmetry a form of planar cell polarity?. *Development* 136:355-66.

Ax, P. 2003. Nemathelminthes. In Multicellular Animals (pp. 1-7). Springer Berlin Heidelberg.

Ax, P. 2012. Multicellular animals: A new approach to the phylogenetic order in nature (Vol. 1). Springer Science & Business Media.

Besson, S. and Dumais, J. 2011. Universal rule for the symmetric division of plant cells. *Proc. Natl. Acad. Sci. USA* 108(15): 6294-6299.

Biggelaar, J.A.M. and Guerrier, P. 1983. Origin of spatial organization. In *The Mollusca: Development*, Vol. 3 (N. H. Verdonk, J. A. M. van den Biggelaar and A. S. Tompa eds.) pp. 179-213. Academic Press, New York.

Boschetti, C., Ricci, C., Sotgia and C., Fascio, U. 2005. The development of a bdelloid egg: a contribution after 100 years. In *Rotifera X* (pp. 323-331). Springer Netherlands.

Boyer, B.C., Henry, J.J. and Martindale, M.Q. 1998. The cell lineage of a polyclad turbellarian embryo reveals close similarity to coelomate spiralians. *Dev. Biol. 204*(1):111-123.

Brusca, R.C. and Brusca, G.J. 1990. Invertebrates (Vol. 2) Sunderland, Massachusetts. Sinauer Associates.

Cameron, R.A. and Davidson, E.H. 1991. Cell type specification during sea urchin development. *Trends. Genet.* 7: 212–218

Cardona, A., Hartenstein, V. and Romero, R. 2005. The embryonic development of the triclad *Schmidtea polychroa. Dev. Gen. Evol. 215*(3):109-131.

Cardona, A., Hartenstein and V. Romero, R. 2006. Early embryogenesis of planaria: a cryptic larva feeding on maternal resources. *Dev. Gen. Evol. 216*(11):667-681.

Cather, J.N. 1963. A time schedule of the meiotic and early mitotic stages of Illyanassa. *Caryologia* 16:663-70.

Cerfontaine, P. 1906. Recherches sur le développement de l'amphioxus. *Arch. Biol. Liege.* 22:229-418. pl. XII–XXII

Chipman, A.D. 2015. "Hexapoda: Comparative Aspects of Early Development."*Evolutionary Developmental Biology of Invertebrates 5*. Springer Vienna, pp. 93-110.

Colwin, A.L. and Colwin, L.H. 1951. Relationships between the egg and larva of Saccoglossus kowalevskii (Enteropneusta): axes and planes: general prospective signifi cance of the early blastomeres. *J. Exp. Zool.* 117:111–137

Costello, D.P. and Henley, C. 1976. Spiralian development: a perspective. *Amer. Zool.* 16: 277-291.

Cowan, C.R. and Hyman, A.A. 2004. Asymmetric cell division in C. elegans: cortical polarity and spindle positioning. *Annu. Rev. Cell. Dev. Biol* 20:427-53.

Damen, W.G. (2007). Evolutionary conservation and divergence of the segmentation process in arthropods. *Dev. Dyn.* 236(6), 1379-1391.

Dan, K. (1960) Cyto-embryology of echinoderms and amphibia. *Int. Rev. Cytol.* 9:321–367

Danilchik, M.V., Brown, E.E. and Riegert, K. 2006. Intrinsic chiral properties of the Xenopus egg cortex: an early indicator of left-right asymmetry?. *Development* 133:4517-26.

Darras, S. and Nishida, H. 2001. The BMP signaling pathway is required together with the FGF pathway for notochord induction in the ascidian embryo. *Development* 128:2629-2638.

Davidson, E.H. 1991. Spatial mechanisms of gene regulation in metazoan embryos. *Development* 113:1-26.

Davidson, E.H., Cameron, R.A. and Ransick, A. 1998 Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. *Development* 125:3269–3290

Day, S. and Lawrence, P.A. 2000. Measuring dimensions: the regulation of size and shape. *Development* 127:2977-87.

de Beauchamp, P. 1929. Le développement des Gastrotriches. *Bull. Soc. Zool. Fr.* 54:549–558

Delile, J., Doursat, R. and Peyriéras, N. 2013. Computational modeling and simulation of animal early embryogenesis with the MecaGen platform. *Comp. Syst. Biol.* 2:359-405.

Donoghue, P.C. and Dong, X.P. 2005. Embryos and ancestors. *Evolving Form and Function: Fossils and Development. Yale Peabody Museum, New Haven*, 81-99.

Dunn, C.W., Giribet,G., Edgecombe, G.D. and Hejnol, A. 2014. Animal phylogeny and its evolutionary implications. *Ann. Rev. Ecol. Evol. Syst. 45*, 371-395.

Driesch, H. and Morgan, T.H. 1895. Zur Analysis der ersten Entwickelungsstadien des Ctenophoreneies. *Arch. Ent. Organ* 2:204–224

Egger, B., Lapraz, F., Tomiczek, B., Müller, S., Dessimoz, C., Girstmair, J., Škunca, N., Rawlinson, K. A., Cameron, C. B., Beli, E., Todaro, M.A., Gammoudi, M., Noreña and C., Telford, M.I. 2015. A Transcriptomic-Phylogenomic Analysis of the Evolutionary Relationships of Flatworms". *Curr. Biol.* 25 (10): 1347–1353.

Ehlers, U. 1985. Phylogenetic relationships within the Plathelminthes. In: The Origins and Relationships of Lower Invertebrates. S Conway Morris, JD George, R Gibson, HM Platt, Eds. Clarendon Press, Oxford. pp 143-158

Eitel, M., Giudu, L., Hadrys, H., Balsamo, M., Schierwater, B. 2011. New insights into placozoan sexual reproduction and development. PloS One 6:e19639.

Elpatiewsky, V.W. 1909. Die Urgeschlechtszellenbildung bei Sagitta . *Anat. Anz.* 35:226–239

Ereskovski, A.V. 2002. Polyaxial cleavage in sponges (Porifera): a new pattern of metazoan cleavage. In *Dokl. Biol. Sci.* 386(1):472-474.

Evans, T., Rosenthal, E.T., Youngblom, J., Distel, D. and Hunt, T. 1983. Cyclin: a protein specified by maternal MRNA in sea urchin eggs that is destroyed at each cleavage division. *Cell* 33(2):389-396.

Freeman, G. and Lundelius, J.W. 1982. The developmental genetics of dextrality and sinistrality in the gastropod Lymnaea peregra. *Roux. Arch. Dev. Biol.* 191:69-83.

Freeman, G. and Lundelius, J.W. 1992. Evolutionary implications of the mode of D quadrant specification in coelomates with spiral cleavage. *J. Evol. Biol. 5*(2), 205-247.

Fritzenwanker, J.H., Genikhovich, G., Kraus, Y. and Technau, U. 2007. Early development and axis specification in the sea anemone Nematostella vectensis. *Dev. Biol*. 310(2):264-279.

García-Varela, M. and Nadler, S.A. 2006. Phylogenetic relationships among Syndermata inferred from nuclear and mitochondrial gene sequences. *Mol. Phyl. Evol. 40*(1), 61-72.

Gazave, E., Lapébie, P., Ereskovsky, A.V., Vacelet, J., Renard, E., Cárdenas, P., and Borchiellini, C. 2012. No longer Demospongiae: Homoscleromorpha formal nomination as a fourth class of Porifera. *Hydrobiologia*, *687*(1), 3-10.

Gilbert, S.F. and Raunio, A.M. 1997. Embryology: Constructing the organism. Sunderland, MA: Sinauer Associates.

Gillies, T.E. and Cabernard C. 2011. Cell division orientation in animals. *Curr. Biol.* 21:R599-609.

Goldstein, B. 1995. Cell contacts orient some cell division axes in the Caenorhabditis elegans embryo. *J. Cell. Biol.* 129:1071-80.

Goldstein, B. 2001. On the evolution of early development in the Nematoda. *Phil. Trans. R. Soc. B Biol. Sci.* 356:1521-1531.

Goulding, M.Q. 2009. Cell lineage of the Ilyanassa embryo: evolutionary acceleration of regional differentiation during early development. *PloS ONE* 4:e5506.

Grande, C. and Patel, N.H. 2008. Nodal signalling is involved in left–right asymmetry in snails. *Nature* 457:1007-11.

Graner, F. and Glazier, J.A. 1992. Simulation of biological cell sorting using a two-dimensional extended Potts model. Phys. Rev. Lett. 69:2013-16.

Grassé, P. P. (1949). Ed. Traite de zoologie. Anatomie, systématique, biologie. Tome VI. Onychophores, Tardigrados, Arthropodes, Trilobltomorphes, Gbélicérates. Ed. Traite de zoologie. Anatomie, systématique, biologie. Tome VI. Onychophores, Tardigrados, Arthropodes, Trilobltomorphes, Chélicérates.

Grassé, P.P. 1975. Treatise on zoology. Anatomy, systematics, biology. Masson.

Grbic, M. 2003. Polyembryony in parasitic wasps: evolution of a novel mode of development. *Int. J. Dev. Biol.* 47:633–642

Green, J.L., Inoue, T. and Sternberg, P.W. 2008. Opposing Wnt pathways orient cell polarity during organogenesis. *Cell* 134:646-56.

Grell, K.G. 1971. Embryonalentwicklung beiTrichoplax adhaerens FE Schulze. *Naturwissenschaften* 58.11: 570-570.

Guerrier, P. 1970. Les caracteres de la segmentation et la determination de la polarite dorsoventrale dans le developpement de quelques Spiralia I. Les formes a premier clivage egal. *J. Embryol. Exp. Morph.* 23:611-37.

Halanych, K.M., Bacheller, J., Liva, S., Aguinaldo, A.A., Hillis, D.M. and Lake, J.A. 1995. 18S rDNA evidence that the Lophophorates are Protostome Animals. *Science* 267: 1641–1643

Hejnol, A. 2010. A twist in time—the evolution of spiral cleavage in the light of animal phylogeny. *Int. Comp. Biol.* 50:695-706.

Henley, C.L. 2012. Possible origins of macroscopic left-right asymmetry in organisms. *J. Stat. Phys.* 148:741-75.

Henry, J.Q. 2014. Spiralian model systems. *Int J Dev Biol* 58:389-401.

Hertzler, P.L. and Wallis, H.C. 1992. Cleavage and gastrulation in the shrimp Sicyonia ingentis: invagination is accompanied by oriented cell division.*Development* 116:127-40.

Hervé, P., Lartillot, N. and Brinkmann, H. 2005. Multigene Analyses of Bilaterian Animals Corroborate the Monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia. *Mol. Biol. Evol.* 22(5): 1246–1253

Hodda, M. 2011. Phylum Nematoda Cobb, 1932. In: Zhang, Z.-Q. (Ed.) Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. *Zootaxa* 3148: 63–95.

Holland, N.D. 1991. Echinodermata: Crinoidea. In: Giese AC, Pearse JS, Pearse VB (eds) Reproduction of marine invertebrates, vol VI, Echinoderms and Lophophorates. Boxwood, Pacifi c Groove

Holland, L.Z. and Yu, J.K. 2004. Cephalochordate (amphioxus) embryos: procurement, culture, basic methods. *Meth. Cell. Biol.* 74:195–215

Holland, L.Z. 2015. Cephalochordata. In: *Evolutionary Developmental Biology of Invertebrates 6* (pp. 91-133). Springer Vienna.

Holton, B., Wedeen, C.J., Astrow, S.H., and Weisblat, D.A. 1994. Localization of polyadenylated RNAs during teloplasm formation and cleavage in leech embryos. *Roux's Arch. Dev. Biol*. 204: 46–53.

Honda, H., Motosugi, N., Nagai, T., Tanemura, M. and Hiiragi, T. 2008. Computer simulation of emerging asymmetry in the mouse blastocyst. *Development* 135:1407-14.

Kajita, A., Yamamura, M. and Kohara, Y. 2003. Computer simulation of the cellular arrangement using physical model in early cleavage of the nematode Caenorhabditis elegans. *Bioinformatics* 19:704-16.

Kanayama, M., Akiyama-Oda, Y. and Oda, H. 2010. Early embryonic development in the spider Achaearanea tepidariorum : microinjection verifi es that cellularization is complete before the blastoderm stage. *Arthrop. Struct. Dev. 39:436–445*

Kaul-Strehlow, S. and Röttinger, E. 2015. Hemichordata. In: *Evolutionary Developmental Biology of Invertebrates 6* (pp. 59-89). Springer Vienna.

Kaufman, Z.S. 2004. On some features of early embryonic development stages of cnidaria. *Russ. J. Mar. Biol.* 30(4):288-292.

Kuroda, R., Endo, B., Abe, M. and Shimizu, M. 2009. Chiral blastomere arrangement dictates zygotic left–right asymmetry pathway in snails. *Nature* 462:790-4.

Lahl, V., Sadler, B. and Schierenberg, E. 2006. Egg development in parthenogenetic nematodes: variations in meiosis and axis formation. *Int. J. Dev. Biol. 50*(4), 393.

Lall, S., Ludwig, M.Z. and Patel, N.H. 2003. Nanos plays a conserved role in axial patterning outside of the Diptera. *Curr. Biol. 13*(3), 224-229.

Lambert, J.D. and Nagy, L.M. 2003. The MAPK cascade in equally cleaving spiralian embryos. *Dev. Biol.* 263:231–41.

Lapraz, F., Rawlinson, K.A., Girstmair, J., Tomiczek, B., Berger, J., Jékely, G. and Egger, B. 2013. Put a tiger in your tank: the polyclad flatworm Maritigrella crozieri as a proposed model for evo-devo. *EvoDevo*, *4*(1), 1.

Lecuit, T. and Lenne, P.F. 2007. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nat. Rev. Mol. Cell. Biol.* 8:633-44.

Lester, S.M. 1988. Ultrastructure of adult gonads and development and structure of the larva of Rhabdopleura normani (Hemichordata: Pterobranchia). *Acta. Zool.* 69:95–109

Leys, S.P. and Ereskovski, A.V. 2006. Embryogenesis and larval differentiation in sponges. *Can. J. Zool.* 84(2):262-287.

Lu, M.S. and Johnston, C.A. 2013. Molecular pathways regulating mitotic spindle orientation in animal cells. *Development* 140:1843-56.

Lynch, J.A., Peel, A.D., Drechsler, A., Averof, M. and Roth, S. 2010 EGF signaling and the origin of axial polarity among the insects. *Curr. Biol.* 20(11):1042–1047

Machida, R. 2006. Evidence from embryology for reconstructing the relationships of hexapod basal clades. Arth. Syst. Phyl. 64:95–104

Malakhov, V.V. 1990. Description of the development of *Ascopodaria discreta*(Coloniales, Barentsiidae) and discussion of the Kamptozoa status in the animal kingdom. *Zool. Zh.* 6920-6930 (In Russian, English summary).

Manoharan, V.N., Elsesser, M.T., and Pine, D.J. 2003. Dense packing and symmetry in small clusters of microspheres. *Science*, *301*(5632), 483-487.

Marin-Riera, M., Brun-Usan, M., Zimm, R., Valikangas, T. and Salazar-Ciudad, I. 2015. Computational modelling of development by epithelia, mesenchyme and their interactions: A unified model. *Bioinformatics*. Btv527.

Martín-Durán, J.M. and Egger, B. 2012. Developmental diversity in free-living flatworms. *EvoDevo*, *3*(1):1.

Maslakova, S.A., Martindale, M.Q. and Norenburg, J.L. 2004. Fundamental properties of the spiralian developmental program are displayed by the basal nemertean Carinoma tremaphoros (Palaeonemerta, Nemertea). *Dev. Biol.* 267:342-360.

Merkel, J., Wollesen, T., Lieb, B. and Wanninger, A. 2012. Spiral cleavage and early embryology of a loxosomatid entoproct and the usefulness of spiralian apical cross patterns for phylogenetic inferences. *BMC Dev. Biol.* 12:11.

Meshcheryakov, V.N. and Beloussov, L.V. 1975. Asymmetrical rotations of blastomeres in early cleavage of gastropoda. *Roux. Arch. Dev. Biol.* 177:193-203.

Meshcheryakov, V.N. 1978. Orientation of cleavage spindles in pulmonate mollusks. I. Role of blastomere shape in orientation of second division spindles. *Ontogenez* 9:558-566.

Minc, N., Burgess, D. and Chang, F. 2011. Influence of cell geometry on division-plane positioning. *Cell* 144:414-426.

Minc, N. and Piel, M. 2012. Predicting division plane position and orientation. *Trends Cell. Biol.* 22:193-200.

Minelli, A. 2010. The origins of larval forms: what the data indicate, and what they don't. *BioEssays*, *32*(1), 5-8.

Morin, X. and Bellaïche, Y. 2011. Mitotic spindle orientation in asymmetric and symmetric cell divisions during animal development. *Dev. Cell.* 21:102-119.

Müller, Y., Ammar, D., and Nazari, E. 2004. Embryonic development of four species of palaemonid prawns (Crustacea, Decapoda): pre-naupliar, naupliar and post-naupliar periods. *Rev. Bras. Zool.* 21(1):27-32.

Müller, W., Hassel, M. and Grealy, M. 2015. *Development and reproduction in humans and animal model species*. Springer.

Munro, E.M. and Odell, G.M. 2002. Polarized basolateral cell motility underlies invagination and convergent extension of the ascidian notochord. *Development* 129:13-24.

Munro, E., Robin, F. and Lemaire, P. 2006. Cellular morphogenesis in ascidians: how to shape a simple tadpole. *Curr. Op. Gen. Dev.* 16:399-405.

Namigai, E.K., Kenny, N.J. and Shimeld, S.M. 2014. Right across the tree of life: The evolution of left–right asymmetry in the Bilateria. *Genesis*, *52*(6), 458-470.

Neves, R.C., Møbjerg Kristensen, R. and Funch, P. 2012. Ultrastructure and morphology of the cycliophoran female. *J. Morph. 273*(8), 850-869.

Newell, G.E. 1948. A contribution to our knowledge of the life history of Arenicola marina L. *J. Mar. Biol. Ass. UK* 27:554-80.

Newman, S.A. 2011. Animal egg as evolutionary innovation: a solution to the 'embryonic horglass puzzle'. *J. Exp. Biol. B Mol. Dev. Evol.* 316:467-83.

Newman, T.J. 2005. Modeling multi-cellular systems using sub-cellular elements. *Math. Biosci. Eng.* 2:611-22.

Newman, S.A. 2011. Animal egg as evolutionary innovation: a solution to the “embryonic hourglass” puzzle. *J. Exp. Zool. B: Mol. Dev. Evol.* 316:467-483.

Newton, W.D. 1970. Gastrulation in the turbellarian Hydrolimax grisea (Platyhelminthes; Plagiostomidae): formation of the epidermal cavity, inversion and epiboly. *Biol. Bull*. 139:539-548.

Nielsen, C. 1994. Larval and Adult Characters in Animal Phytogeny. *Am. Zool.* 34:492-501.

Nielsen, C. 2002. The phylogenetic position of Entoprocta, Ectoprocta, Phoronida, and Brachiopoda. *Int. Comp. Biol. 42*(3): 685-691.

Nielsen, C. 2012. *Animal evolution: interrelationships of the living phyla*. Oxford University Press on Demand.

Okazaki, K. 1975. Spicule formation by isolated micromeres of the sea urchin embryo. *Am. Zool.* 15:567–581

Ortolani, G. 1963. Origine dell’organo apicale e dei derivati mesodermici nello sviluppo embrionale di Ctenofori. *Acta. Embryol. Morphol. Exp.* 7:191–200

Patel, N.H. 2000. It's a bug's life. *Proc. Natl. Acad. Sci. USA. 97*(9), 4442-4444.

Pawlak, J.B., Sellars, M.J., Wood, A. and Hertzler, P.L. 2010. Cleavage and gastrulation in the Kuruma shrimp Penaeus (Marsupenaeus) japonicus (Bate): a revised cell lineage and identification of a presumptive germ cell marker. *Dev. Growth Differ. 52*(8): 677-692.

Perseke, M., Hankeln, T., Weich, B., Fritzsch, G., Stadler, P.F., Israelsson, O., Bernhard, D. and Schlegel, M. 2007. The mitochondrial DNA of Xenoturbella bocki: genomic architecture and

phylogenetic analysis. *Theor. Biosc.* 126(1): 35-42

Price, A.L. and Patel, N.H. 2008. Investigating divergent mechanisms of mesoderm development in arthropods: the expression of Ph-twist and Ph-mef2 in Parhyale hawaiensis. *J. Exp. Zool. B: Mol. Dev. Evol. 310*(1):24-40.

Raff, R.A. and Byrne, M. 2006. The active evolutionary lives of echinoderm larvae. *Heredity* 97:244–252

Raven, C.P. 1967. The distribution of special cytoplasmic differentiations of the egg during early cleavage in Limnaea stagnalis. *Dev. Biol.* 16:407-37.

Rawlinson, K.A. 2010. Embryonic and post-embryonic development of the polyclad flatworm Maritigrella crozieri; implications for the evolution of spiralian life history traits. *Front. Zool.* 7(1): 1.

Riegler, A. 2008. Natural or internal selection? the case of canalization in complex evolutionary systems. *Artif. Life.* 14(3):345-362.

Robert, H. 1902. Embryologie des Troques. *Arch. de Zool. Exper. et Gen.* 3:18-538.

Rogulja, D., Rauskolb, C. and Irvine, K.D. 2008. Morphogen control of wing growth through the Fat siganling pathway. *Dev. Cell.* 15:309-21.

Rouse, G.W., Wilson, N.G., Carvajal, J.I. and Vrijenhoek, R.C. 2016. New deep-sea species of Xenoturbella and the position of Xenacoelomorpha. *Nature* 530(7588):94-97.

Ruiz-Trillo, I., Riutort, M., Littlewood, D.T.J., Herniou, E.A. and Baguñà, J. 1999. Acoel flatworms: earliest extant bilaterian metazoans, not members of Platyhelminthes. *Science* 283(5409):1919-1923.

Rupert, E.E. and Barnes, R.D. 1994. Invertebrate Zoology. Saunders Coll.Publ., Fort Worth, Philadeplphia, etc.

Rupert, E. E., Fox, R.S. and Barnes, R.D. 2004. Invertebrate zoology: a functional evolutionary approach. Brooks/Cole, Belmont, CA.

Sacks, M. 1955. Observations on the embryology of an aquatic gastrotrich, Lepidodermella squamata (Dujardin, 1841). *J. Morphol.* 96:473–495

Salazar-Ciudad, I., Solé, R.V., and Newman, S.A. 2001. Phenotypic and dynamical transitions in model genetic networks II. Application to the evolution of segmentation mechanisms. *Evol. Dev. 3*(2):95-103.

Salazar-Ciudad, I., Jernvall, J. and Newman, S.A. 2003. Mechanisms of pattern formation in development and evolution. *Development* 130:2027-2037.

Salazar-Ciudad, I. 2010. Morphological evolution and embryonic developmental diversity in metazoa. *Development* 137:531–539

Sander, K. 1976. Specifi cation of the basic body pattern in insect embryogenesis. *Adv. Insect. Physiol.* 12:125–238

Sandersius, S.A. and Newman, T.J. 2008. Modeling cell rheology with the subcellular element model. *Phys. Biol.* 5:015002.

Satoh, N., Rokhsar, D. and Nishikawa, T. 2014. Chordate evolution and the three-phylum system. *Proc. R. Soc. B.* 281(1794):20141729.

Schierenberg, E. and Junkersdorf, B. 1992. The role of eggshell and underlying vitelline membrane for normal pattern formation in the early C. elegans embryo. *Roux's Arch. Dev. Biol. 202*(1):10-16.

Schierenberg, E. 2001. Three sons of fortune: early embryogenesis, evolution and ecology of nematodes. *Bioessays*, 23(9):841-847.

Schierenberg, E. 2006. Embryological variation during nematode development, T. E. Community, Ed. Wormbook. [Online]. Available: http://www.wormbook.org

Schmidt, G. 1985. Development and life cycles. In: Crompton D, Nickol B (eds) Biology of the Acanthocephala. Cambridge University Press, Cambridge, pp. 273–305

Schmidt-Rahesa, A. 1999. Nematomorpha. In: Knobil, E, Neill, JD. Eds. Encyclopedia of reproduction. New York Academic Press, 333-340.

Schmidt-Rhaesa, A. 2007. Evolution of organ systems. Oxford University Press.

Scholtz, G., Ponomarenko, E. and Wolff, C. 2009. Cirripede cleavage patterns and the origin of the Rhizocephala (Crustacea: Thecostraca). *Arthr. Syst. Phyl.* 67:219-228.

Scholtz, G. and Wolff, C. 2013. Arthropod embryology: cleavage and germ band development. In: Arthropod Biology and Evolution. Springer Berlin Heidelberg, pp 63-89.

Seydoux, G. and Schedl, T. 2001. The germline in C. elegans: origins, proliferation, and silencing. *Int. Rev. Cytol.* 203:139–185

Shimotori, T. and Goto, T. 2001. Developmental fates of the first four blastomeres of the chaetognath Paraspadella gotoi: relationship to protostomes. *Dev. Growth Differ.* 43(4):371-382.

Siewing, R. 1979. Homology of cleavage-types? *Fortschr. Zool. Syst. Evolutionsf* 1:7–18

Slack, J. 2014. Establishment of spatial pattern. *Wiley. Interdiscip. Rev. Dev. Biol.* 3:379-88.

Smith, J.M., Cridge, A.G. and Dearden, P.K. 2010. Germ cell specification and ovary structure in the rotifer Brachionus plicatilis. *EvoDevo*, *1*(1): 1.

Song, C., Wang, P. and Makse, H.A. 2008. A phase diagram for jammed matter. *Nature* 453(7195):629-632.

Spek, J. 1926. Über gesetzmässige Substanzverteilung bei der Furchung des Ctenophoreneis und ihre Beziehung du dem Determinationsproblem. *Arch. Entwicklungsmech. Org.* 107:54–73

Srivastava, M., Begovic, E., Chapman, J., Putnam, N.H., Hellsten, U., et al. 2008. The Trichoplax

genome and the nature of placozoans. *Nature* 454:955-919.

Stevens ,N.M. 1904. On the germ cells and the embryology of Planaria simplicissima. *Proc. Acad. Nat. Sci. Philadelphia.* 56:208-220.

Strathmann, M.F. 1987. Reproduction and development of marine invertebrates of the northern Pacific coast: data and methods for the study of eggs, embryos, and larvae. University of Washington Press.

Strome, S., Garvin, C., Paulsen ,J., Capowski, E., Martin, P., and Beanan, M. 1994. Specification and development of the germline in Caenorhabditis elegans. *Ciba Found. Symp.* 182, 31–45

Summers, R.G., Stricker, S.A. and Cameron, R.A. 1993. Applications of confocal microscopy to studies of sea urchin embryogenesis. *Methods. Cell. Biol.* 38:265–287

Świderski, Z., Poddubnaya, L.G., Gibson, D.I., Levron, C., and Młocicki, D. 2011. Egg formation and the early embryonic development of Aspidogaster limacoides Diesing, 1835 (Aspidogastrea: Aspidogastridae), with comments on their phylogenetic significance. *Parasitol. Int. 60*(4):371-380.

Tagawa, K., Nishino, A., Humphreys, T., Satoh, N. 1998. The spawning and early development of the Hawaiian acorn worm (hemichordate), Ptychodera flava . *Zool. Sci*. 15:85–91

Takahashi, H. and Shimizu, T. 1997. Role of intercellular contacts in generating an asymmetric mitotic apparatus in the Tubifex embryo. *Dev. Growth. Differ.* 39:351-62.

Tassy, O., Dauga, D., Daian, F., Sobral, D., Robin, F., Khoueiry, P., et al. 2010. The ANISEED database: digital representation, formalization, and elucidation of a chordate developmental program. *Gen. Res.* 20(10):1459-1468.

Technau, U., Genikhovich, G., and Kraus, J. E. 2015. Cnidaria. In: Evolutionary Developmental Biology of Invertebrates 1 .pp. 115-163. Springer Vienna.

Tee, Y.H., Shemesh, T., Thiagarajan, V., Hariadi, R.F., Anderson, K.L., Page, C., Volkmann, N., Hanein, D., Sivaramakrishnan, S., Kozlov, M.M., Bershadsky, A.D. 2015. Cellular chirality arising from the self-organization of the actin cytoskeleton. *Nat. Cell. Biol*. 17:445-57.

Telford, M.J., Bourlat, S.J., Economou, A., Papillon, D. and Rota-Stabelli, O. 2008. The evolution of the Ecdysozoa. *Phil. Trans. Roy. Soc. B: Biol. Sci.* 363(1496): 1529-1537.

Telford, M.J., Budd, G.E. and Philippe, H. 2015. Phylogenomic insights into animal evolution. *Curr. Biol.* 25(19):876-887.

Théry, M. and Bornens, M. 2006. Cell shape and cell division. *Curr. Opin. Cell. Biol.* 18:648-57.

Timothy, D., Littlewood, J., Telford, M.J. and Bray, R.A. 2004. Protostomes and Platyhelminthes. In Cracraft, J. & Donoghue, M.J. *Assembling the tree of life*. Oxford University Press US. pp. 209–223

Uglem, G.L. and Larson, O.R. 1969. The life history and larval development of Neoechinorhynchus saginatus Van Cleave and Bangham, 1949 (Acanthocephala: Neoechinorhynchidae). *J. Parasitol.* 1212-1217.

Valentine, J.W. 1997. Cleavage patterns and the topology of the metazoan tree of life. *Proc. Natl. Acad. Sci. USA. 94*(15):8001-8005.

Wallace, R.L., Ricci, C. and Melone, G. 1996. A cladistic analysis of pseudocoelomate (Aschelminth) morphology. *Invert. Biol.* 115: 104-112.

Wandelt, J. and Nagy, L.M. 2004. Left-right asymmetry: more than one way to coil a shell. *Curr. Biol*. 14:R654-6.

Weisblat, D.A. and Huang, F.Z. 2001. An overview of glossiphoniid leech development. *Can. J. Zoo*l. 79(2):218-232.

Willey, C.H. and Godman, G.C. 1951. Gametogenesis, fertilization and cleavage in the trematode, Zygocotyle lunata (Paramphistomidae). *J. Parasitol. 37*(3):283-296.

Williamson, D.I. 2009. Caterpillars evolved from onychophorans by hybridogenesis. *Proc. Natl. Acad. Sci. USA. 106*(47):19901-19905.

Wray, G.A. 2000. The evolution of embryonic patterning mechanisms in animals. In *Seminars in cell & developmental biology* 11(6): 385-393. Academic Press.

Wang, S.W., Griffin, F.J. and Clark, W.H.Jr. 1997. Cell-cell association directed mitotic spindle orientation in the early development of the marine shrimp Sicyonia ingentis. *Development* 124:773-80.

Wilson, E.B. 1892. The cell lineage of Nereis. A contribution to the cytogeny of the annelid body. *J. Morph.* 6:361-480.

Wilson, E.B. 1893. Amphioxus, and the mosaic theory of development. *J. Morphol.* 8:579-639

Wolff, C. and Scholtz, G. 2002. Cell lineage, axis formation, and the origin of germ layers in the amphipod crustacean Orchestia cavimana. *Dev. Biol.* 250:44–58

Wu, H.R., Wu, H.R., Chen, Y.T., Su, Y.H., Luo, Y.J., Holland, L.Z. and Yu, J.K. 2011. Asymmetric localization of germline markers Vasa and Nanos during early development in the amphioxus Branchiostoma fl oridae. *Dev. Biol.* 353:147–159

Yin, L., Zhu, M., Knoll, A. H., Yuan, X., Zhang, J. and Hu, J. 2007. Doushantuo embryos preserved inside diapause egg cysts. *Nature*, 446(7136), 661-663.

Younossi-Hartenstein, A. and Hartenstein, V. 2000. The embryonic development of the polyclad flatworm Imogine mcgrathi. *Dev. Gen. Evol. 210*(8-9):383-398.

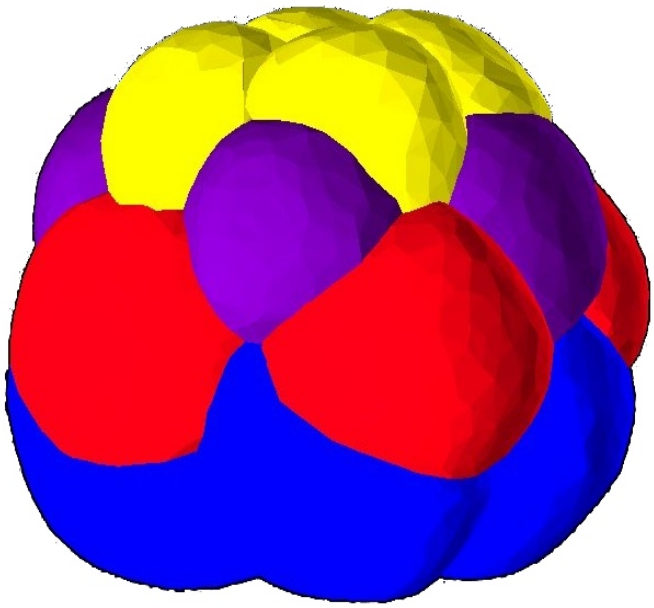
Zhang, S.C., Holland, N.D. and Holland, L.Z. 1997. Topographic changes in nascent and early mesoderm in amphioxus embryos studied by DiI labeling and by in situ hybridization for a Brachyury gene. *Dev. Gen. Evol.* 206:532-535

Zhang, Z. 2013. Animal biodiversity: An update of classification and diversity in 2013. In: Zhang, Z.-Q. (Ed.) Animal Biodiversity: An Outline of Higher-level Classification and Survey of Taxonomic Richness. *Zootaxa* 3703(1): 5–11

CHAPTER II

Simple cell processes lead to spiral cleavage and to its variation

The content of this chapter is a scientific paper which has been accepted (September 2016) in the journal *Development*, so this is currenly awaiting publication.



Simple cell processes lead to spiral cleavage and to its variation

Running title: Modelling spiral cleavage

Miguel Brun-Usan[a,b], Miquel Marín-Riera[a,b], Cristina Grande[c], Marta Truchado-Garcia[c], Isaac Salazar-Ciudad[a,b*]

[a] Genomics, Bioinformatics and Evolution. Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Barcelona, Spain.
[b] Evo-devo Helsinki community, Centre of excellence in computational and experimental developmental biology, Institute of Biotechnology, University of Helsinki, PO Box 56, FIN-00014 Helsinki, Finland.
[c] Departamento de Biología Molecular and Centro de Biología Molecular, “Severo Ochoa” (CSIC, Universidad Autónoma de Madrid), Madrid, Spain

*-Author for correspondence: isalazar@mappi.helsinki.fi

Keywords: Spiral cleavage, developmental rules, developmental morphospace

ABSTRACT

During cleavage the zygote becomes partitioned into a set of cells with a particular spatial arrangement. Spiral cleavage is the most abundant cleavage type at the phylum level. Different cellular processes have been hypothesized to be responsible for the development of the specific spatial arrangement of blastomeres in the spiral blastula. These include the orientation of cell division according to an animal-vegetal gradient, according to cells' main axis (Hertwig's rule), according to the contact areas between cells or orthogonally to previous divisions (Sach's rule). Cell adhesion and cortical rotation have also been proposed to be involved in spiral cleavage.

We use a computational model of cell and tissue bio-mechanics to implement the different existing hypotheses about how the specific spatial arrangement of cells in spiral cleavage arises during development. We found that cell polarization by an animal-vegetal gradient, Sachs' rule, cortical rotation and adhesion, when combined, reproduce the spiral cleavage while other combinationsof processes can not. We reproduce the cell spatial arrangement of the blastulae of seven different species (four snails, two polychaetes and a nemertean).

INTRODUCTION

Most metazoans start their development through a series of fast cell divisions that partition the zygote into a set of blastomeres. In this cleavage process a specific spatial cell arrangement (what we call in here the cleavage pattern) arises in each species. There are several types of cleavage in metazoa (Gilbert and Raunio, 1997). The spiralian cleavage is the most abundant cleavage type at the phylum level. It is found in mollusks, annelids and nemerteans. Other lophotrochozoan phyla (platyhelminthes, rotifers, brachiopods, phoronids, gastrotrichs, and bryozoans) also exhibit spiral cleavage in at least some of their species (Hejnol, 2010). In spite of having a very similar cleavage these phyla have very different adult morphologies. The ensemble of phyla with spiralian cleavage has been suggested to be a monophyletic group (Nielsen, 1994; Laumer et al., 2015), the *Spiralia*.

As in other types of cleavage, spiralian cleavage begins with two successive nearly meridional cell divisions giving rise to four large cells (termed A, B, C and D macromeres) that lie in a plane perpendicular to the animal-vegetal (A-V) axis of the egg. These four macromeres then divide towards the animal pole, giving rise to four, usually smaller, cells (micromeres). In contrast with the radial cleavage found in many other metazoans, in spiral cleavage each micromere is not placed right over its sister macromere, but is displaced to the right (or to the left depending on the

organism) of its sister macromere. The third cell division, thus, proceeds at an oblique angle relative to the A-V axis. In this process all micromeres are displaced in the same direction (either all to the right or all to the left). If the third cell division produces a micromere to the right, the next cell division will proceed to the left, and the ensuing ones will follow in a right-left alternation (the reverse alternation occurs if the third division is to the left). This alternation makes that, when viewed from the animal pole, the new micromeres seem to spin clockwise or counterclockwise when they arise (Gilbert and Raunio, 1997; Henry, 2014). Compared with those of other groups, spiralian embryos tend to undergo fewer cell divisions before gastrulation, making it easier to follow the fate of blastomeres. When cell fates are compared between different spiralian, it appears that the same adult or larval organs in different species arise from the same blastomeres (defined by lineage and relative position in the blastula (Nielsen, 1994)). The first four cell division rounds are synchronous but this synchrony is gradually lost over developmental time and the similarity between groups becomes less obvious.

There is also abundant literature focusing in the signaling events that, taking place within a specific spatial blastomere arrangement, lay out the cell fate of each blastomere (Freeman and Lundelius, 1982; Lambert and Nagy, 2003; Kuroda et al., 2009; Grande and Patel, 2008). However, much less is known about how the specific spatial arrangement of blastomeres in spiral cleavage is attained (although there is some work in the early morphogenesis of some invertebrate non-spiralian models, most notably in ascidians (Munro et al., 2006)). There are a number of developmental processes that have been hypothesized to explain spiral cleavage. Some are roughly understood as developmental rules by which the direction of the cell division plane is determined during cleavage (Freeman and Lundelius, 1982) while others are processes of cell mechanical interaction that lead to cell displacement during cleavage (Meshcheryakov and Beloussov, 1975; Wandelt and Nagy, 2004; Henley, 2012). In here, we call each of the former group of hypotheses *developmental rules of division plane specification* or, simply, rules, while we reserve the term *non-directional processes* for the hypothesis not related to the direction of the division plane. Our aim here is to assess which of these previously proposed rules, either alone or in combination, are capable of producing the spiral pattern when implemented in a realistic bio-mechanical model of cleavage. These rules are (see Fig. 1 and Chapter 3 of this Thesis for details):

1) Hertwig's rule (Fig. 1A): In many developmental systems cells tend to divide with their division plane perpendicular to the cell's longest axis (Minc et al., 2011; Minc and Piel, 2012). It has been proposed that the Hertwig's rule could explain part of the spiral cleavage pattern by translating blastomeres' shape changes (due to the packing produced by cell-cell adhesion) into specific

directions of the cell division planes (Meshcheryakov, 1978).

2) Cell polarization rule (Fig. 1A): Cells can be polarized and this polarization can determine the direction in which cells divide. The direction of this polarization can be affected by clues in the cell's surroundings or by asymmetries in the oocyte cytosol inherited by the blastomeres (see SI). Cells, then, tend to divide perpendicularly to the direction of cell polarization. This can happen either because cell growth is biased towards the direction of polarization (Rogulja et al., 2008) (and then the Hertwig's rule leads to cells dividing perpendicularly to the direction of polarization) or because polarization directly affects the direction of division (Morin and Bellaïche, 2011). The latter case is more relevant here since there is no cell growth during cleavage. It has been suggested that cell polarization in blastomeres could arise because of a gradient in the distribution of some molecules in the cell cortex (Freeman and Lundelius, 1982; Lu and Johnston, 2013). Such an asymmetry would, by promoting a differential attachment of the astral microtubules to the part of the cortex with higher concentration of these molecules, regulate the tilting of the mitotic apparatus relative to the A-V axis prior to cytokinesis and then bias cell divisions to take place along a specific direction (Freeman and Lundelius, 1982; Lu and Johnston, 2013).

3) Cell-Cell contact rule (Fig. 1A): The direction of division in a cell is also known to be affected by which parts of it are in contact with other cells. When this rule applies, cells would tend to divide towards (or against in some cases (Wang et al., 1997)) the part of the cell contacting adjacent cells. This has been suggested to occur because adhesion in a cell region would modify the underlying cell cortex so that astral microtubules are stabilized in this region, increasing their traction of the mitotic spindle (Hertzler and Wallis, 1992; Goldstein, 1995; Théry and Bornens, 2006). This rule has been proposed to have a role in the cleavage of some spiralian species (e.g. *Tubifex* worms (Takahasi and Shimizu, 1997)).

4) Sachs' rule (Fig. 1A): During spiral cleavage each cell division after the first three division rounds tends to be perpendicular to the previous cell division. Concerning the spiral cleavage, this has been suggested to be crucial to explain the left-right alternation of cell divisions after the four-cell stage (Guerrier, 1970; Meshcheryakov and Beloussov, 1975; Henley, 2012). This rule has been proposed to arise from the stereotypic duplication of the centrioles (that form a 90º angle between them) between cell divisions, that, in turn, biases the position of the mitotic spindle towards perpendicularity (Théry and Bornens, 2006; Minc and Piel, 2012).

The non-directional processes (those not affecting the direction of the division plane but that may affect spiral cleavage) are:

5) Cortical rotation (Fig. 1B): According to Meshcheryakov's and related works (Meshcheryakov and Beloussov, 1975; Wandelt and Nagy, 2004; Henley, 2012), blastomeres in spiral embryos rotate over themselves just after cell division. This rotation occurs in the same sense (clock-wise or counter-clockwise around the rotation axis depending on the embryo) in all blastomeres. These authors argue that the left/right twist between 4 and 8-cell stage and the relative placement of blastomeres is produced by this rotation (See SI). This rotation has been observed in many developmental systems (Meshcheryakov and Beloussov, 1975; Danilchik et al., 2006), but its potential role for cleavage remains unclear (Henley, 2012). This mechanism does not specify the direction of cell division *per se*, but it may affect the mechanical interactions between cells after division and lead to changes in the relative positioning of the blastomeres.

6) Cell-cell adhesion (Fig. 1C): Cell adhesion increases the contact surface between cells (the greater the adhesion strength the larger and flatter the contact surface between two cells) (Lecuit and Lenne, 2007; Sandersius and Newman, 2008). This way adhesion can lead to cell shape changes that may then affect blastomere shape, which blastomeres are in closer contact, the direction of cell division (e.g. through Hertwig's rule) and, overall, the spatial arrangement of cells in the blastula.

7) Asymmetric cell division (Fig. 1D): Cell division can give rise to daughter cells of different size. Several mechanisms could produce that. In many embryos, intra-cellular gradients of molecules (e.g. PAR proteins in *C. elegans* (Cowan and Hyman, 2004)) regulate the relative size of daughter cells (Gilbert and Raunio, 1997; Munro et al., 2006; Salazar-Ciudad et al., 2003). These gradients promote asymmetric cell division by a differential binding of the spindle microtubules to the cortex that produces asymmetric pulling forces and a displacement of the contractile ring during cytokinesis (Gilbert and Raunio, 1997; Lu and Johnston, 2013).

This rule may play a role in generating morphological variation in the spiral cleavage patterns because blastomeres of different relative size are placed and compacted in different manners due to cell-cell adhesion (Merkel et al., 2012).

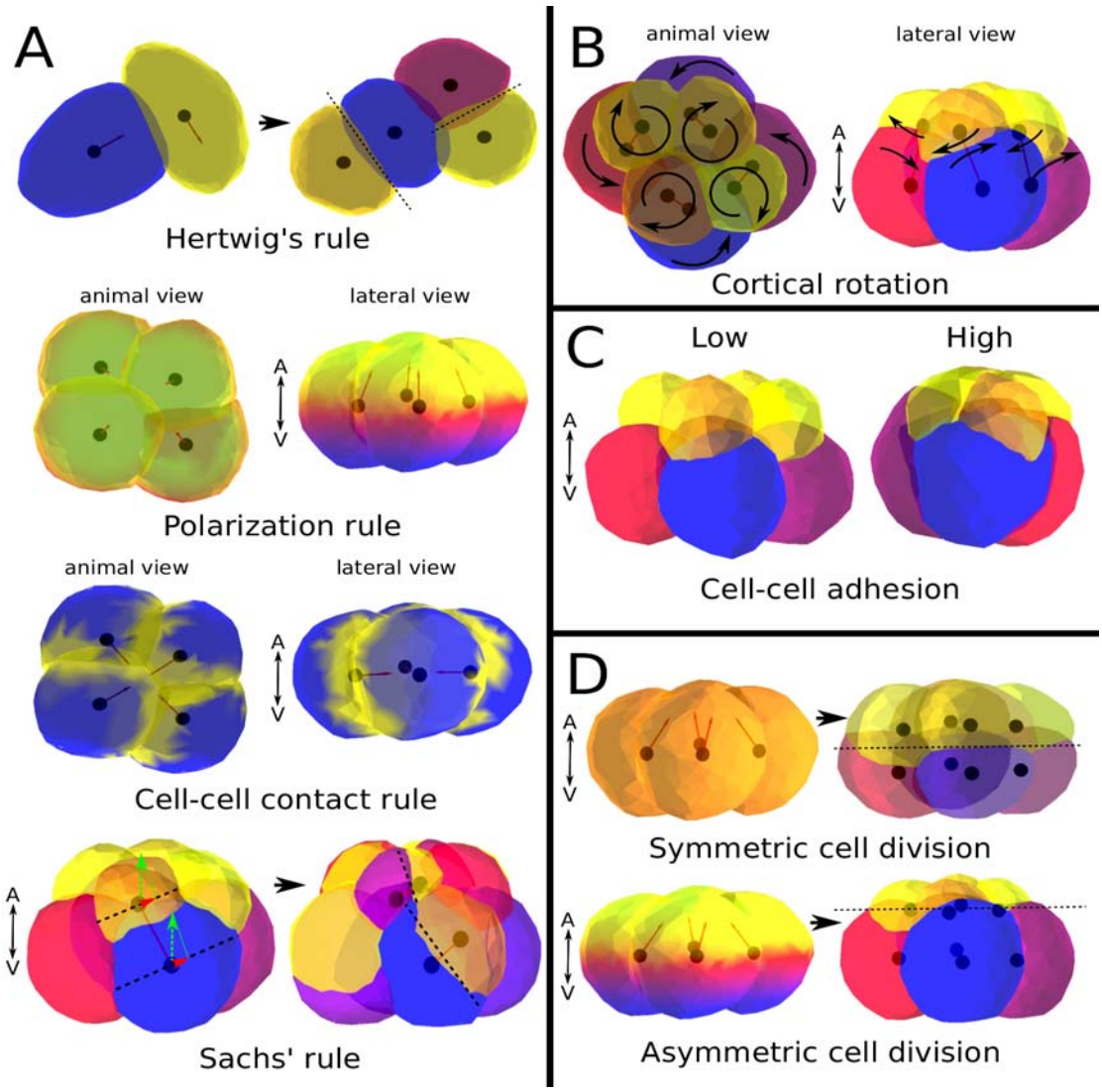


Figure 1. Basic depiction of the developmental rules and non-directional processes considered in this work. In all panels, black dots represent cell's centroids. A) Different rules that cells use to specify the direction of cell division (small arrows). From top to down: Hertwig's rule makes cells to divide perpendicular to the cell's longest axis. Polarization rule: cells divide perpendicular to the direction of a molecular gradient along the A-V axis (yellow parts of cells have the highest concentration of that molecule and blue parts the least). Cell-cell contact rule: cells divide perpendicular to the areas of contact of a blastomere with other with other blastomeres (yellow parts of cells). Sachs' rule: cells divide at right angles to the previous cell division. The direction of cell division (green arrows) specified by one of the developmental direction rules are projected in a plane (dashed lines) that is perpendicular to the previous cell division. The direction of cell division (red arrowheads) is then perpendicular to the previous one and to another acting developmental direction rule. B) Cortical rotation. Blastomeres rotate over themselves just after cell division (black arrows) around a rotation axis that is defined when it arises from its mother cell (dark lines linking cell's centroids). C) Cell-cell adhesion increases the surface of contact between cells (the greater the adhesion strength the larger and flatter the contact surface between two cells, and the more rounded the blastula shape). D) Asymmetric cell division. Intra-cellular gradients of molecules regulate the relative size of daughter cells. Colors represent molecule concentration in each part of the cells. The steeper the gradient the more different the size of daughter cells.

Notice that, by combining these rules, we look for the minimal (sufficient) set of mechanisms capable of generating the spiral pattern, although more complex (and redundant) mechanisms can also exist. For instance, recent work in *Lymnaea* snails (Shibazaki et al., 2004) shows that, during the transition from 4 to 8-cell stage, the clockwise cleavage pattern (that gives rise to adults with dextral coiling) exhibits cytoskeletal processes not observed in the counterclockwise pattern. If a counterclockwise spiral cleavage can be produced without such additional processes, it means that they are not strictly necessary for spiral cleavage, but are superimposed to a more basic (by default) mechanisms (Wandelt and Nagy, 2004; Henley, 2012). Such a basic set of mechanisms is what we refer as a minimal set, irrespective of whether they produce clockwise or counterclockwise spiral pattern.

We use a mathematical model, the SpiralMaker, in which each cell is implemented as a set of subcellular elements (spheric elastic volumes) in three-dimensional space. The physical interactions between the subcellular elements (hereafter nodes), make each whole cell to display visco-elastic properties alike to those observed in real cells (Newman, 2005; Sandersius and Newman, 2008; Marin-Riera et al., 2015). Thus, nodes can adhere to each other (preferentially to nodes from the same cells but also to nodes from other cells) but repel each other if they are too close (implementing the physical fact that two cells can not occupy exactly the same position in space). These attraction and repulsion forces, together with some noise, lead to node movement and, consequently, to changes in cell shape and spatial location within the blastula. The model also implements cell behaviors such as cell division, polarization and cell adhesion.

All our simulations start after the second blastomere division (before this stage, the spiral cleavage proceeds as in other, non-spiral cleavage types). Thus, in all our simulations the initial conditions consist of four blastomeres (See Fig. 1 in Chapter 3). In summary, in each blastula simulation we only specify the initial conditions, which rules are used and their relative strength. On the contrary the 3D spatial position and shape of each blastomere over time is an output of the model resulting from the model dynamics.

On the basis of the above proposed developmental rules we generate a theoretical developmental morphospace of possible spiral cleavage patterns. Each axis of this morphospace corresponds to one of the rules implemented in the model and along each of those dimensions the relative contribution of the respective rule is quantitatively varied. Thus, the direction of the cell division is then simply the weighted sum of each of the 3D vectors specifying cell division direction according to each rule in a given cell.

We explore all combinations of three rules. In addition all simulations include cell adhesion. We also ran some simulations with an external coverage or eggshell (see SI). Although this

coverage is not required for the spiral cleavage to occur (Freeman and Lundelius, 1982; Kuroda et al, 2009), it can affect the blastomere arrangement by limiting the available space for them (Kajita et al., 2003). The distribution of cleavage patterns within the morphospace reflects how the different developmental rules give rise to different cleavage patterns. This allows us to evaluate in detail the cleavage pattern forming capabilities of each rule and of the combination of different rules in a multi-cellular context (see methods).

Since nodes in the model can contain molecules it is possible to implement intracellular molecular gradients. These gradients regulate some of the above explained rules (most notably cell polarization and asymmetric division: see the methods for a more detailed description). Non-directional processes such as adhesion, have parameters of their own (e.g. specifying the adhesion strength between cells) (see SI).

We also compare the blastulae in our simulated morphospace with those of several species to see how well these rules could reproduce the spiral cleavage patterns of concrete species. We focused on species that were most available experimentally to us or for which information about the cleavage pattern has been published already. Those species are the gastropods *Crepidula*, *Planorbella*, *Lottia* and *Trochus* (one species for each main snail groups), the nemertean worm (ribbon worm) *Carinoma* and the polychaete worms *Nereis* and *Arenicola*. For the three first gastropod species we obtained 16-cell stage embryos and stained them to measure the volume of each blastomere. The relative volumes of these blastomeres were then compared with those of blastulae in the simulated morphospace. For the other species we used only data from published work (Wilson, 1892; Robert, 1902; Newell, 1948; Freeman and Lundelius, 1982; Maslakova et al., 2004; Goulding, 2009). Since individual blastomere volumes were not published we used, instead, the list of blastomeres that are in physical contact with each other blastomere (that was present in the publications). These relative contacts between specific blastomeres, that are known to be crucial for the inductive mechanisms acting on later development (Lambert and Nagy, 2003; Grande and Patel, 2008; Kuroda et al, 2009), were then compared between empirical data and simulations.

There are a number of previous models on cleavage. Some focus on the bio-mechanical properties (rheology) of blastula-like aggregates and compared their results qualitatively with mammalian blastulae (Honda et al., 2008; Sandersius and Newman, 2008). Some others try, as here, to disentangle the morphogenetic processes responsible for the different types of cleavage existing in animals. The model that is most similar to ours is (Kajita et al., 2003) but it applies only to the first two divisions of *Caenorhabditis elegans* (a non-spiralian) and does not include the entire set of rules we include in here. Another recent model tries to explain the radial cleavage of the sea urchin (Akiyama et al., 2010). This latter model, however, is only 2D and it includes only three of the

developmental rules we consider in here.

RESULTS

1. Minimal set of rules accounting for the spiral cleavage pattern:

Fig. 2A shows the only combination of rules and rule parameters in which the spiral pattern arises. This combination is the one that includes the cell polarization rule, the Sachs' rule, cortical rotation and inter-cellular adhesion (see volume encircled in red in the parameter space, Fig. 2A). Other combinations of three rules are also be plotted (see Fig. 6 in Chapter 3) but they do not produce any cleavage pattern resembling the spiral one. A cleavage pattern was considered to be spiral only if: 1) the blastomeres were organized in groups of four cells (quartets) along the A-V axis, forming the four blastomeres of each quartet a square in a plane perpendicular to the A-V axis; 2) the blastomeres closer to the animal pole were in close contact between them; 3) sister blastomeres were obliquely positioned in respect to the A-V axis and 4) this oblique positioning was in the same sense for blastomeres dividing at the same time (either all to the right or all to the left of their sister blastomere).

The cell polarization rule in our model, if not combined with other rules, leads cell division to be oriented towards the animal pole (this is up and and towards the embryo's center; Fig. 1A). The results of our model show that this rule is strictly required for spiralian cleavage to proceed normally after the third division, although on its ow it can not lead to spiral cleavage. With this rule the four new micromeres arising in the third cell division (4 to 8-cell stage) are close to each other, enabling mechanical interaction between them that, as we explain below, is important for further spiral cleavage.

Adhesion tends to increase the contact surface between blastomeres. Because of that each new cell, in the model, tends to gradually place itself, after division, between two other existing cells (in the case of the third division these latter cells are the macromeres).This leads, in the model, to a relative displacement, a twist, of each micromere in respect to its sister macromere in the resulting 8-cell stage. If no other rules apply this displacement is random in direction (on average 50% of the times to the left and 50% to the right). As a result not always does each micromere end up correctly positioned between two macromeres since the twist between two adjacent micromeres may be in contrary directions and then both of them may try to position themselves between the same two macromeres (a feature not found in spiral patterns in nature, Fig. 6A in Chapter 3). In fact, in the 4-cell stage of Spiralia all cell divisions are directed either to the right or to the left. It is clear

then that something in addition to the adhesion and polarization rules is required.

Our results show that, in the 8-cell stage, cortical rotation, and no other rule, leads all micromeres to be displaced in the same direction in respect to their sister macromeres. With the cortical rotation each blastomere in the 8-cell stage rolls over its neighbors in the same sense and places itself correctly between a distinct pair of underlying macromeres. The displacement of the micromeres occurs in all of them at the same time and we found that this requires adhesion between micromeres and macromeres and between micromeres. In the former the nodes in the micromeres tend to adhere more effectively to nodes from the macromeres and as a result each micromere tends to maximize its contact surface with macromeres. If the cell adhesion in the model is high enough this maximization of contact surface can only be achieved if each micromere is placed between two macromeres. Adhesion between micromeres, instead, produces a friction-like effect in the contact areas between rotating micromeres, so that nodes in these contact areas move less than nodes elsewhere in the cell. As a result, the effective rotation of the cell surface introduced by the rotation rule is stronger in those areas than elsewhere in the cell. This produces a coherent displacement of the four micromeres in the same clockwise or counterclockwise direction (Figs. 3 in Chapter 3). Without adhesion the rotation rule simply causes each micromere to rotate slightly around its own axis without any net displacement or blastula-level change in morphology (Fig. 3 in Chapter 3).

Adhesion and cortical rotation together can not account for the spatial distribution of blastomeres in the spiral pattern after the 8-cell stage. With these rules, blastomeres always divide towards the animal pole and form cone shaped blastulae (Fig. 6 in Chapter 3). However, our results indicate that the spiral pattern can be produced beyond the 8-cell stage if the Sachs' rule is also included (Fig. 2A, encircled). With this rule the fourth division round occurs at an angle from the A-V axis. This angle is not exactly 90 degrees because, as described, the actual division direction is also affected by the polarization rule.

In most of our simulations each combination of the rules and processes is applied during the whole developmental time of each embryo. To further explore the relative roles of each rule and non-directional process we also performed simulations in which some of the rules were interrupted after some stage. This analysis shows that cortical rotation is crucial between the 4 and 8-cell stage to position the micromeres between the macromeres forming an oblique angle in respect to the A-V axis. The Sachs' rule leads divisions in successive rounds to be perpendicular between one another and then the right oblique orientation in the 8 cells stage translates into an alternation between right and left oblique cell divisions from that stage onwards (although modified also by the polarization rule). Thus, from the 8-cell stage onwards cortical rotation is not required, being the polarization, Sachs' and adhesion rules sufficient for the spiral pattern to arise.

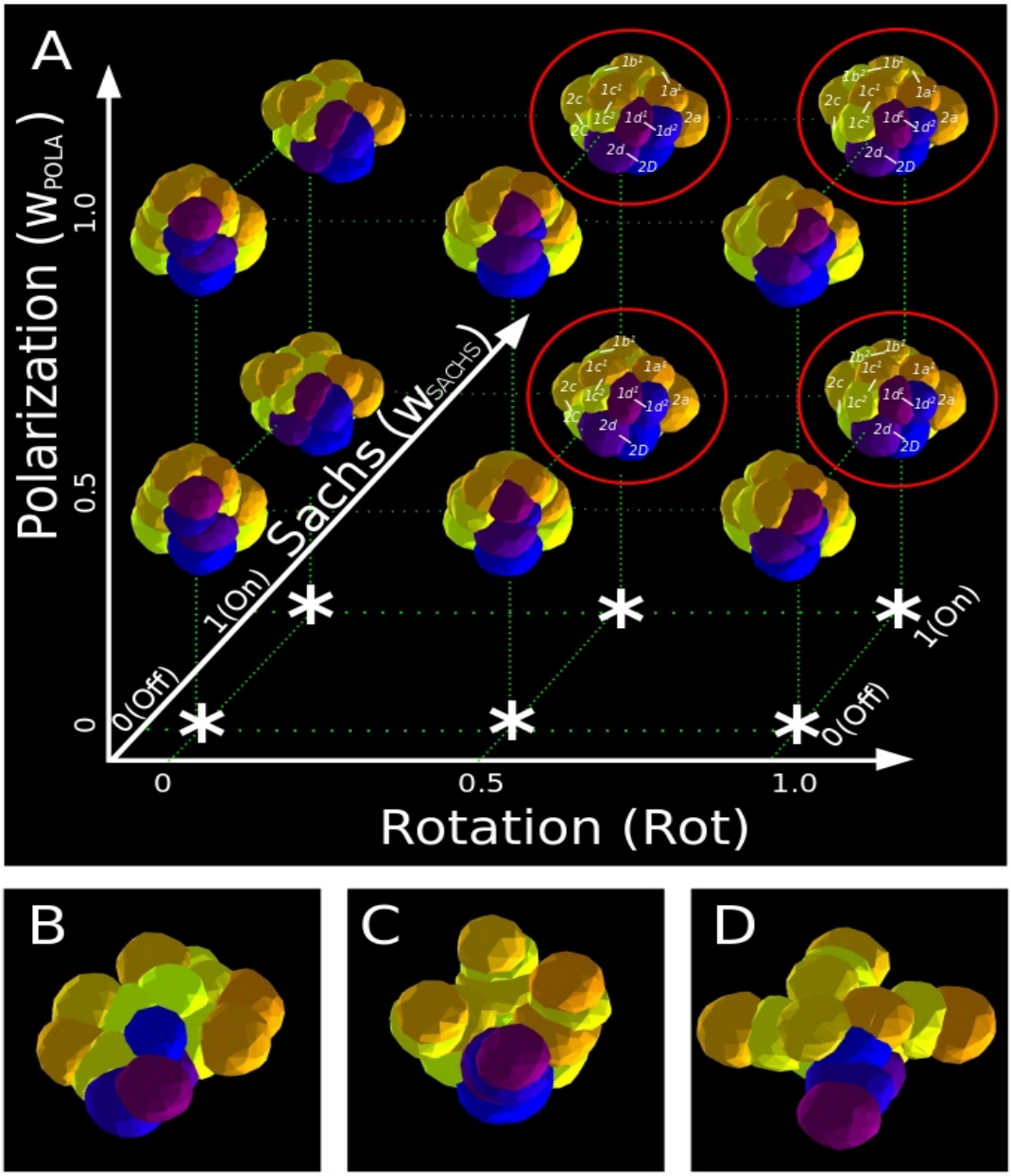


Figure 2. A specific combination of basic developmental rules reproduces the spiral pattern. A) The combination of rule and rule parameters that can produce spiral cleavage patterns (encircled in red). The spiral cleavage pattern occurs when the polarization rule is applied along with Sachs' rule, cortical rotation and adhesion. Asterisks indicate hypervolumes of the morphospace that are not definable (where the combination of rules does not unambiguously specify the direction of cell division). All blastulae are drawn from a slightly lateral animal view, with the animal pole to the top. B-D Some examples of non-spiral cleavage patterns arising from other combinations of rules: B) Hertwig and cell-cell contact rule, low adhesion; C) Low adhesion, no Sachs, cell polarization and cell-cell contact rule and D) cell-cell contact rule, moderate rotation and low adhesion. Blastulae in lateral animal view.

2. Failure of the other rules to lead to spiral cleavage:

As Figs. 2B and Fig. 5 of Chapter 3 show, many cleavage patterns from the 8-cell stage onwards (60% of all combinations of rules) lack the tetra-radial symmetry along the A-V axis that is characteristic of many cleavage patterns in animals (spiralians or not). Those asymmetric cleavage patterns are found if Hertwig's rule is applied. According to this rule the direction of division is parallel to the longest cell axis. Since blastomeres tend to be spherically shaped, and in our model they are made of a finite number of nodes, small fluctuations around this spherical shape can completely change in which direction the longest cell axis is and, thus, the direction of division. The direction of division is then quite sensitive to noise (note that noise is present also in real systems). Thus, when this rule is applied, the direction of cell division in the third round of cell division can be different in each of the four macromeres (albeit in general, each micromere appears in the "upper" hemisphere of its respective macromere). As a result the four new micromeres may not lie at the same relative position along the A-V axis. These relatively small misalignments between macro- and micromeres at the 8-cell stage are amplified during further cell divisions, leading to blastulae with irregular cleavage patterns.

The spiral pattern is not found in the simulations including the cell-cell contact rule (neither alone or combined with other rules). Our simulations show that with this rule the direction of cell division tends to point towards the center of the blastula (Fig. 6C in Chapter 3). Since no two cells can occupy the same physical space (in the model cells that are very close repel each other) no extra cells can be placed in the center of the blastula. As a result, daughter cells are passively displaced towards the periphery of the blastula over developmental time. In the resulting cleavage patterns, micromeres are not densely packed and are arranged in a radial manner that minimizes, or even removes the contacts between them (Fig. 2C, 2D). This lack of contact between adjacent blastomeres diminishes the effect of cortical rotation because blastomeres can not roll over their adjacent neighbors (nor over their sister blastomeres). When the cell-cell contact rule is combined with the polarization rule, new micromeres appear towards the animal pole, but they tend to avoid the center of the blastula (when viewed from the animal pole) so that they form a sort of cavity between them (Fig. 2C). The resulting open configuration prevents the emergence of spiral pattern, in which the four animal-most micromeres are always in contact with one another.

3. Cleavage patterns of seven different spiralian species:

To further explore the accuracy of our model we systematically varied the parameters (increments of 20% from their minimum to their maximum possible values) in each of the rules that we found to lead to spiral patterns and compared the resulting variation in spiral patterns with those of seven different spiralian species. In those simulations asymmetric cell division was also implemented, making that in each cell division daughter cells would have different sizes according to their position in a gradient along the A-V axis (see methods and SI). That way we obtained a large set of simulated blastulae which were compared with the blastulae of some spiralian species. Asymmetric cell division was included in here but not in the analysis above because in the seven species simulated cells are of different size along the A-V axis but our simulations above show that asymmetric division *per se* is not required for achieving the spiral pattern itself.

For the gastropods *Crepidula fornicata, Planorbella duryi* and *Lottia gigantea* we first obtained 16-cell stage blastulae and then measured the relative volumes of each blastomere (relative to the embryo volume; see SI).

Within our simulated morphospace we encountered blastulae closely resembling those of these three species. In those simulated blastulae the volume of each blastomere is very similar to that of the corresponding blastomere in one of these three species. By corresponding blastomere we mean blastomeres with the same cell lineage in the simulations and in the empirically measured species (Conklin nomenclature, Gilbert and Raunio, 1997) (see Fig. 3A). According to the model the subtle differences in the spiral patterns of these three species arise from differences in the asymmetry of cell division, in adhesion and cortical rotation (see SI for the exact numerical differences). For *Planorbella* and *Lottia*, the visual resemblance between the simulated and the real blastulae improved when an eggshell with moderate compressive effect was included in the simulations (see Fig. 3 and Chapter 3). In the case of *Crepidula*, although the volumes of the real and simulated blastomeres were almost identical, small discrepancies in the blastomere arrangement appear. These discrepancies could be attributed to variations in the mechanical properties (e.g. surface tension) between blastomeres of very disparate size that are not implemented in the model.

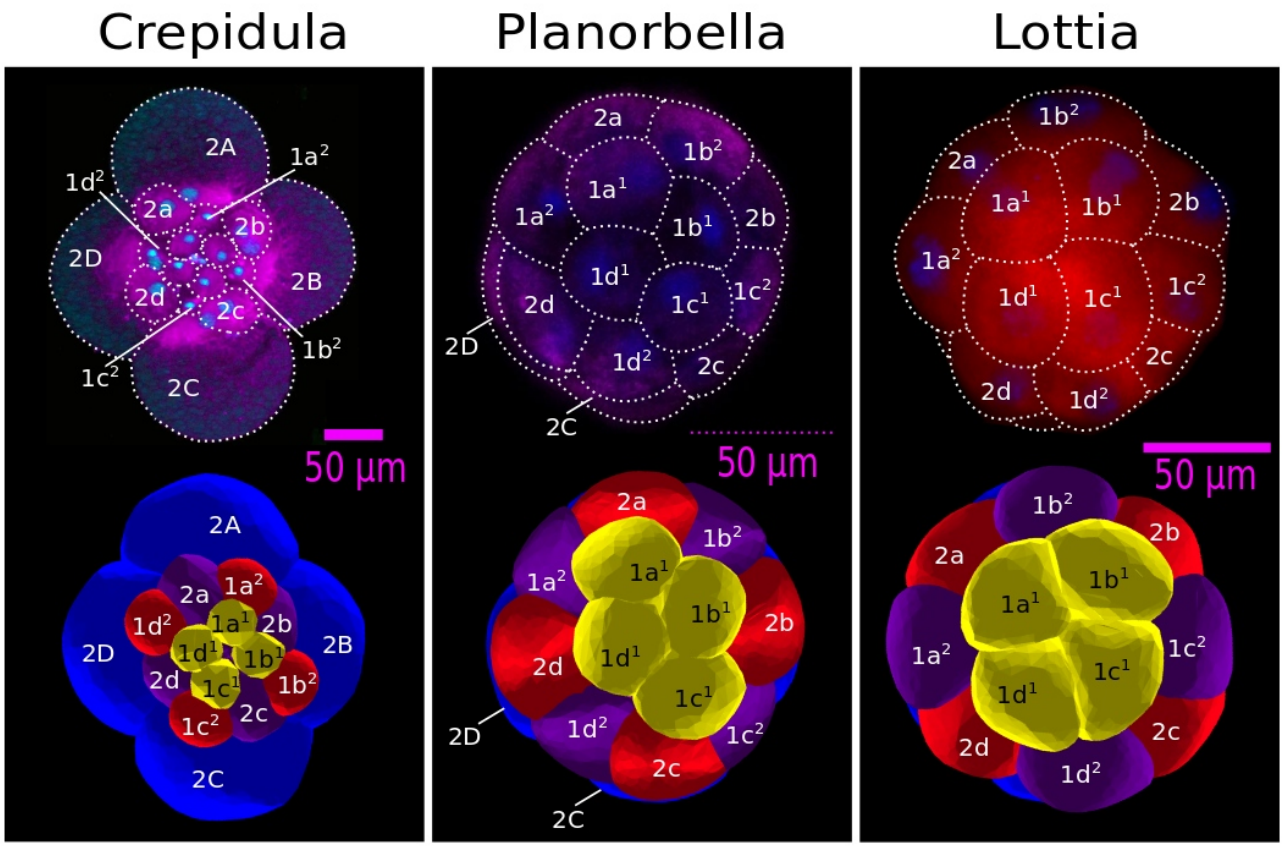


Figure 3. Cleavage patterns reminiscent of these of three spiralian species (gastropod mollusks) found by comparing the relative volumes between the different blastomeres. Upper row: animal views of the 16-cell stage real embryos stained with anti-ß-tubulin E7 (from which the relative volumes of each blastomere were measured). Lower row: animal views of the simulated cleavage patterns best matching each species blastula. Similarity is measured as the sum of the square of the differences in the relative volumes of the blastomeres between real and simulated embryos. Blastomeres are labeled according to Conklin nomenclature. In the simulated blastulae, colours represent the cells' generation (which is also reflected by the blastomeres' labels). Simulated cleavage patterns of *Planorbella* and *Lottia* are surrounded by an eggshell (not shown). The small discrepancies between real and simulated *Crepidula* patterns could be attributed to variations in the mechanical properties (e.g. surface tension) between blastomeres of very disparate size that are not implemented in the model.

The spiral patterns of *Trochus niloticus*, *Carinoma tremaphoros*, *Nereis diversicolor* and *Arenicola marina* were compared with the ones found in the model based on the descriptions present in the literature (Wilson, 1892; Robert, 1902; Newell, 1948; Freeman and Lundelius, 1982; Maslakova et al., 2004; Goulding, 2009). From these descriptions it is possible to get the overall topology of connections between blastomeres, this is which blastomere (defined by cell lineage) are in physical contact with each other (but not the blastomere volumes) (see methods and SI). For

Arenicola and *Nereis*, the simulations used initial conditions in which some the cells, the D-blastomere, were bigger than others (as reported for such species Wilson, 1892; Newell, 1948; Freeman and Lundelius, 1982). As Fig. 4B-E shows, we found simulated cleavage patterns resembling those of these four different spiralian species (see SI for a detailed description of the parameter values that produce each species blastula). This similarity is of at least 90%, this is 90% of the contacts between blastomeres are identical (see Chapter 3 for details). Furthermore, our simulations reveal that the hypervolumes of the parameter space occupied by the different species differ in size and shape (Fig. 7 in Chapter 3).

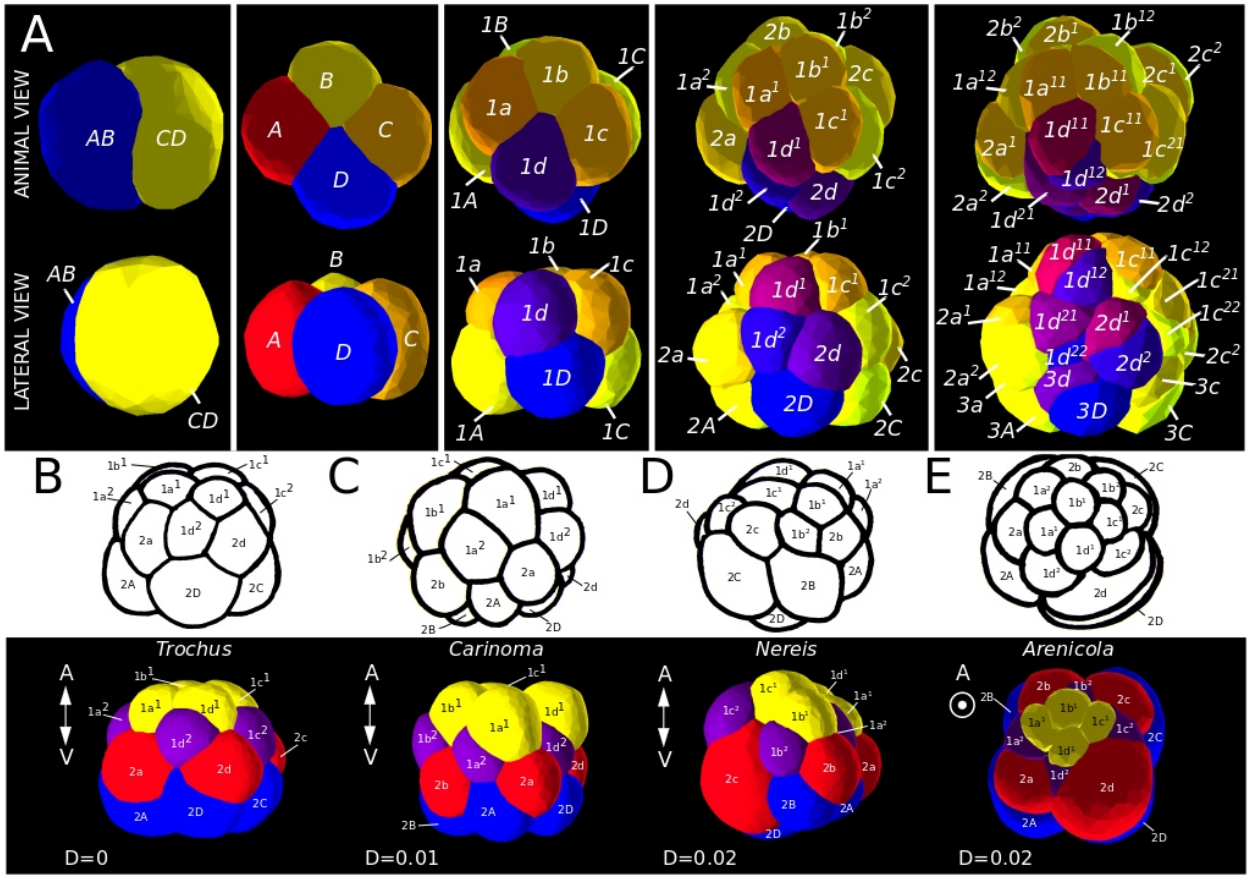


Figure 4. Cleavage patterns of four spiralian species found by comparing the relative contacts between adjacent blastomeres. A) The combination of rules drawn from experiment 1 accurately reproduces the blastomere arrangement found in spiralians until the 32-cell stage. Blastomeres are labeled according to Conklin nomenclature (1). Blastomeres descending from the D-macromere are colored in blue tones. B-E) The four simulated (lower panel) blastulae that best match the real ones (upper panels, extracted from bibliography) for the four species (See Results). Similarity is measured as the proportion of contacts between blastomeres that are the same between real and simulated embryos. Blastomeres are labeled according to Conklin nomenclature. In the simulated blastulae, colours represent the cells' generation (which is also reflected by the blastomeres' labels).

DISCUSSION

Our results indicate that the spatial arrangement of cells characteristic of the spiral cleavage arises mainly due to the combination of an animal-vegetal polarization of cell division, Sachs' rule, cortical rotation and adhesion. Cortical rotation is important in the third division to produce a coherent twist or rotation of all micromeres respect to the macromeres but not later on.

Our modeling of realistic bio-mechanics and rules in the context of 3D blastulae has allowed us to discard other rules and processes that have been suggested to be responsible for spiral cleavage (Meshcheryakov, 1978; Takahasi and Shimizu, 1997; Aw and Levin, 2009). Our approach also suggests that up to the 32-cell stage the formation of the cleavage pattern does not require signaling between cells. A molecular animal-vegetal gradient is, however, required but such kind of gradients are known to be intrinsic to many metazoan oocytes (Raven, 1967; Gilbert and Raunio, 1997; Slack, 2014). This does not imply that early signaling is not important, it simply suggests that cell signaling does not need to be involved in the determination of the direction of cell division to lead to the spiral pattern. This does not imply that signaling is not involved in it, but simply that it is not a logical necessity from what we know about other different cellular processes that could be involved in spiral cleavage. In addition, it is clear that cell signaling may be involved in cell fate determination. These conclusions can not be extended beyond the 32-cell stage. After the 32-cell stage cell divisions are no longer synchronous and processes more complex than the ones considered in our model may start to be important.

Our second experiment shows that part of the morphological variation observed in spiralian cleavage can be explained by quantitative variations in these few rules (this is changes in the relative strength of each of such rules). Species specific patterns, as those of the seven different spiralian species considered in here, arise for particular parameter combinations. Thus, in spite of its simplicity, our model seems to be indicative about the developmental processes underlying morphological differences at the level of blastulae between species. However, in some cases the simulated and real blastulae do not look visually identical. This is likely be due to the intrinsically provisional or imperfect nature of our model, and in most other models. It could be that other cellular processes other than we want we include in this model could explain the 5% or 10% difference between model and reality that we observe. Moreover, the available data (the proportion of identical contacts between specific blastomeres and the similarity between relative blastomere volumes) only capture two aspects (thought very important ones) of the general embryo morphology. Nevertheless, with this model we can already assess if the few developmental rules involved in spiralian cleavage can generate a variation comparable to the one found in nature and

how the general properties of this variation are.

The distribution of the different species analyzed in the parameter space should be informative about which cleavage patterns ought to be easier to produce from other ones by changes in the underlying developmental parameters (Alberch, 1982; Salazar-Ciudad, 2007). We found that some cleavage patterns, (e.g. patterns reminiscent of *Trochus*), can be produced by many different combinations of parameter values (they are found in a large volume of the parameter space) while others require a much more restricted combination of parameter values (e.g. patterns reminiscent of *Arenicola*, 24% of the parameter space leads to this species). We also find that the cleavage patterns of species that are phylogenetically closer, (e.g. *Arenicola* and *Nereis*), are not closer to each other in the parameter space than to the patterns of other, more phylogenetically distant, species. This suggests that the underlying developmental parameters are relatively easy to change (presumably a small number of mutational changes are required) (Newman, 2011). This also suggests that, at least within spiralians and based on our tiny sample, these differences should not be used for phylogenetic inferences at the low level of species or genera (Merkel et al., 2012).

MATERIALS AND METHODS:

The mathematical model used in this study, the SpiralMaker, is a modification of a general model of animal embryonic development, the EmbryoMaker (Marin-Riera et al., 2015), which allows us to implement each of the developmental rules studied in this article. In the EmbryoMaker cells are made of subcellular elements (nodes), which can be conceptualized as spherical portions of a cell occupying a physical volume. Nodes adhere to each other if they touch but repel each other if their centers become to close. A node of a cell adheres with higher affinity to nodes from the same cell that to nodes from different cells (this ensures that cells keep their physical integrity). Motion is computed in continuous time and 3D space by solving a system of differential equations assuming Langevin over-damped dynamics and some degree of noise (see SI). As a result of cell division, adhesion, repulsion and noise the positions of nodes (and thus the position and shape of blastomeres) change over simulation time. This 3D spatial distribution of those nodes represents the embryo's morphology and within each cell it represents cell shape. Notice, thus, that cell shape and position within the embryo, and its variation, are not a free parameter of the model, but results from the model biomechanics. In turn, changes in the spatial location and shape of cells configure the overall changes in an embryo's morphology. Mechanical interactions between nodes, such as the relative cell movement, occur faster than whole-cell processes such as cell division. In the model this is ensured by choosing long time intervals between cell division rounds.

Since in most spiralians cleavage takes place without cell growth, the number of nodes in the SpiralMaker is kept constant during each simulation. Thus the number of nodes in a cell halves (for symmetric cell division) in each cell division round while the number of cells doubles. The SpiralMaker includes only non-epithelial cells, as blastomeres, and only the cell behaviors of cell division, cell polarization and cell adhesion (other types of cells and cell behaviours are included in the EmbryoMaker).

The model includes: 1) an initial condition of the 4-cell stage in which cells have an intracellular molecular gradient along the A-V axis (presumably inherited from the zygote), 2) Cells dividing synchronously at regular time intervals, 3) Using different rules to determine the direction of cell division, 4) Using different slopes of the intracellular gradient affecting the asymmetry of cell division, 5) Using different degrees of cortical rotation, 6) Using different values of adhesion between nodes. All those things are specific of the SpiralMaker but not of the EmbryoMaker from which it derives.

The developmental rules and non-directional processes of the SpiralMaker are implemented in the model as simple rules acting on the cells' elements (See Chapter 3. In an initial experiment (in silico experiment 1, see Chapter 3), we explored whether the combination of these rules and processes can generate the spiral pattern or not.

Each simulation used three different rules (plus adhesion), and different weights in the relative importance of each rule in determining the direction of cell division. Other, high dimensional combinations of rules (eg. four or five rules by simulation) were not considered. We discarded this possibility because our goal (the emergence of spiral cleavage pattern) was achieved under the simpler combination of only three rules (plus adhesion). For this initial exploration the number of nodes in the blastula were 1250, cell division was kept approximately symmetric, and all simulations were run until the 16-cell stage.

In a more detailed approach (in silico experiment 2, see SI) we assess how variations within the spiral pattern (including the cleavage patterns similar to those of several invertebrate species) can be produced. For that, we took all the combinations of rules that in experiment 1 were found to produce spiral patterns and performed a more exhaustive sampling of our theoretical developmental morphospace in each rule and non-directional process parameter (increments of 20% in each parameter from its minimum possible value to its maximum possible value). For this experiment, the number of nodes per simulation was 2500, and simulations were run until the 32-cell stage. In addition, asymmetric cell division was included, and we also considered asymmetry in the size of the blastomeres in initial conditions. We also ran some simulations in which an outer cover (or eggshell) surrounds the blastula (see SI).

ACKNOWLEDGEMENTS:

We thank R. Zimm, I. Salvador-Martinez and P. Hagolani for comments. We also thank the Hybridome bank at Iowa University for developing the mouse anti-ß-tubulin E7 antibody used in this study and the Marine Biological Laboratory (especially Drs. Nipam Patel and Lee Niswander) for supporting the collection and fixation of *C. fornicata* embryos.

COMPETING INTERESTS

The authors declare no competing or financial interests.

AUTHOR CONTRIBUTIONS

M.B.U. and I.S.C. conceived the experiments and wrote the manuscript. M.B.U. performed the experiments. M.B.U., M.M.R. and I.S.C. equally contributed to modelling and data analysis. C.G. and M.T.G. processed and measured the real embryos.

FUNDING

This research was funded by the Finnish Academy (WBS 1250271) and the Spanish Ministry of Science and Innovation (BFU2010-17044) to I.S.C., the Spanish Ministry of Science and Innovation (BES2011-046641) to M.B.U., the Generalitat de Catalunya (2013FI-B00439) to M.M.R., the Spanish MICINN (ref. 2012-052214) to MTG and the UAM and the Spanish MICINN (CGL2011-29916) to CG.

REFERENCES

Akiyama, M., Tero, A. and Kobayashi, R. (2010). A mathematical model of cleavage. *J. Theor. Biol.* 264, 84-94.

Alberch, P. (1982) Developmental constraints in evolutionary processes. In *Evolution and development* (ed. J.T. Bonner), p313–332 .Dahlem Konferenzen. Springer-Verlag: Heidelberg.

Aw, S. and Levin, M. (2009). Is left-right asymmetry a form of planar cell polarity?. *Development* 136, 355-366.

Cowan, C.R. and Hyman, A.A. (2004). Asymmetric cell division in C. elegans: cortical polarity and spindle positioning. *Annu. Rev. Cell Dev. Biol.* 20, 427-453.

Danilchik, M.V., Brown, E.E. and Riegert, K. (2006). Intrinsic chiral properties of the Xenopus egg cortex: an early indicator of left-right asymmetry?. *Development* 133, 4517-4526.

Freeman, G. and Lundelius, J.W. (1982). The developmental genetics of dextrality and sinistrality in the gastropod Lymnaea peregra. *Roux. Arch. Dev. Biol.* 191, 69-83.

Gilbert, S.F. and Raunio, A.M. (1997). Embryology: Constructing the organism. Sunderland, MA: Sinauer Associates.

Goldstein, B. (1995) Cell contacts orient some cell division axes in the Caenorhabditis elegans embryo. *J. Cell. Biol.* 129, 1071-1080.

Grande, C. and Patel, N.H. (2008). Nodal signalling is involved in left–right asymmetry in snails. *Nature* 457, 1007-1011.

Guerrier, P. (1970). Les caracteres de la segmentation et la determination de la polarite dorsoventrale dans le developpement de quelques Spiralia I. Les formes a premier clivage egal. *J. Embryol. Exp. Morph.* 23, 611-637.

Hejnol, A. (2010). A twist in time- the evolution of spiral cleavage in the light of animal phylogeny. *Integr. Comp. Biol.* 50, 695-706.

Henley, C.L. (2012) Possible origins of macroscopic left-right asymmetry in organisms. *J. Stat. Phys.* 148, 741-775.

Henry, J.Q. (2014) Spiralian model systems. *Int. J. Dev. Biol.* 58, 389-401.

Hertzler, P.L. and Wallis, H.C. (1992). Cleavage and gastrulation in the shrimp Sicyonia ingentis:

invagination is accompanied by oriented cell division. *Development* 116, 127-140.

Honda, H., Motosugi, N., Nagai, T., Tanemura, M. and Hiiragi, T. (2008). Computer simulation of emerging asymmetry in the mouse blastocyst. *Development* 135, 1407-1014.

Kajita, A., Yamamura, M. and Kohara, Y. (2003). Computer simulation of the cellular arrangement using physical model in early cleavage of the nematode Caenorhabditis elegans. *Bioinformatics*. 19, 704-716.

Kuroda, R., Endo, B., Abe, M. and Shimizu, M. (2009) Chiral blastomere arrangement dictates zygotic left–right asymmetry pathway in snails. *Nature* 462, 790-794.

Lambert, J.D. and Nagy, L.M. (2003). The MAPK cascade in equally cleaving spiralian embryos. *Dev. Biol.* 263, 231-241.

Laumer, C.E., Bekkouche, N., Kerbl, A., Goetz, F., Neves, R.C., Sørensen, M.V., Kristensen, R.M., Hejnol, A., Dunn, C.W., Giribet, G. et al. (2015). Spiralian phylogeny informs the evolution of microscopic lineages. *Curr. Biol. 25*, 2000-2006.

Lecuit, T. and Lenne, P.F. (2007). Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nat. Rev. Mol. Cell. Biol.* 8, 633-644.

Lu, M.S. and Johnston, C.A. (2013). Molecular pathways regulating mitotic spindle orientation in animal cells. *Development* 140, 1843-1856.

Marin-Riera, M., Brun-Usan, M., Zimm, R., Valikangas, T. and Salazar-Ciudad, I. (2015). Computational modelling of development by epithelia, mesenchyme and their interactions: A unified model. *Bioinformatics*. Btv527.

Merkel, J., Wollesen, T., Lieb, B. and Wanninger, A. (2012). Spiral cleavage and early embryology of a loxosomatid entoproct and the usefulness of spiralian apical cross patterns for phylogenetic inferences. *BMC Dev. Biol.* 12, 11.

Meshcheryakov, V.N. (1978). Orientation of cleavage spindles in pulmonate mollusks. I. Role of

blastomere shape in orientation of second division spindles. *Ontogenez* 9, 558-566.

Meshcheryakov, V.N. and Beloussov, L.V. (1975). Asymmetrical rotations of blastomeres in early cleavage of gastropoda. *Roux. Arch. Dev. Biol.* 177, 193-203.

Minc, N., Burgess, D. and Chang, F. (2011). Influence of cell geometry on division-plane positioning. *Cell* 144, 414-426.

Minc, N. and Piel, M. (2012). Predicting division plane position and orientation. *Trends. Cell. Biol.* 22, 193-200.

Morin, X. and Bellaïche, Y. (2011). Mitotic spindle orientation in asymmetric and symmetric cell divisions during animal development. *Dev. Cell.* 21, 102-119.

Munro, E., Robin, F. and Lemaire, P. (2006). Cellular morphogenesis in ascidians: how to shape a simple tadpole. *Curr. Op. Gen. Dev.* 16, 399-405.

Newman, S.A. (2011). Animal egg as evolutionary innovation: a solution to the 'embryonic horglass puzzle'. *J. Exp. Biol. B. Mol. Dev. Evol.* 316, 467-483.

Newman, T.J. (2005). Modeling multi-cellular systems using sub-cellular elements. *Math. Biosci. Eng.* 2, 611-622.

Nielsen, C. (1994). Larval and Adult Characters in Animal Phytogeny. *Am. Zool.* 34, 492-501.

Raven, C.P. (1967). The distribution of special cytoplasmic differentiations of the egg during early cleavage in Limnaea stagnalis. *Dev. Biol.* 16, 407-437.

Ren, X. and Weisblat, D.A. (2006). Asymmetrization of first cleavage by transient disassembly of one spindle pole aster in the leech Helobdella robusta. *Dev. Biol.* 292, 103-115.

Rogulja, D., Rauskolb, C. and Irvine, K.D. (2008) Morphogen control of wing growth through the Fat siganling pathway. *Dev. Cell.* 15, 309-321.

Salazar-Ciudad, I. (2007). On the origins of morphological variation, canalization, robustness, and evolvability. *Integr. Comp. Biol.* 47, 390-400.

Salazar-Ciudad, I., Jernvall, J. and Newman, S.A. (2003). Mechanisms of pattern formation in development and evolution. *Development* 130, 2027-2037.

Sandersius, S.A. and Newman, T.J. (2008) Modeling cell rheology with the subcellular element model. *Phys. Biol.* 5, 015002.

Shibazaki, Y., Shimizu, M. and Kuroda, R. (2004). Body handedness is directed by genetically determined cytoskeletal dynamics in the early embryo. *Curr. Biol.* 14, 1462-1467.

Slack, J. (2014). Establishment of spatial pattern. *Wiley Interdiscip. Rev. Dev. Biol.* 3, 379-388.

Takahashi, H. and Shimizu, T. (1997). Role of intercellular contacts in generating an asymmetric mitotic apparatus in the Tubifex embryo. *Dev. Growth Differ.* 39, 351-362.

Théry, M. and Bornens, M. (2006). Cell shape and cell division. *Curr. Opin. Cell. Biol.* 18, 648-657.

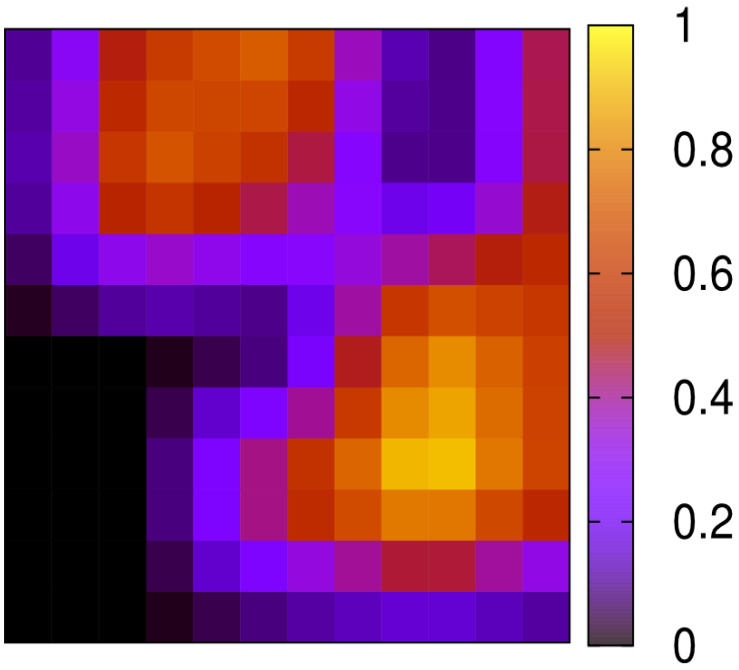
Wandelt, J. and Nagy, L.M. (2004). Left-right asymmetry: more than one way to coil a shell. *Curr. Biol.* 14, R654-656.

Wang, S.W., Griffin, F.J. and Clark, W.H. Jr. (1997). Cell-cell association directed mitotic spindle orientation in the early development of the marine shrimp Sicyonia ingentis. *Development* 124, 773-780.

CHAPTER III

The SpiralMaker model of cleavage.

The content of this chapter constitutes the supplementary material of the scientific paper described in the previous chapter. It explains the inner workings of the SpiralMaker model. That is, how the different cell processes are implemented in this mathematical model.



Contents of this chapter:

A. SpiralMaker

1. Introduction:

The mathematical model used in this work is a modification of the EmbryoMaker: a general modeling framework of animal embryonic development capable of simulating complex developmental processes in 3D (Marin-Riera et al., 2015). We refer to this modification as the "SpiralMaker" model. The SpiralMaker is fully described in this supplementary information without needing reference to EmbryoMaker itself.

EmbryoMaker considers cells, cell parts, gene products and other molecules involved in gene regulation. The model includes in a unified framework bio-mechanical and gene product interactions in development. The model implements epithelial cells, non-epithelial cells and extra-cellular matrix, and also includes all basic cell behaviors of animal cells (*sensu* Salazar-Ciudad et al., 2003). These are: cell growth (polar and non-polar), cell division (directed and non-directed), apoptosis, secretion of ECM and signaling molecules, reception of extracellular signals, cell contraction, cell adhesion and movement and shape change as a consequence of those. By tuning these GRNs and choosing the appropriate initial conditions, EmbryoMaker can be readily applied to simulate different developing organs. In this work, we apply this framework to simulate the spiral cleavage pattern. In order to do so, we only consider non-epithelial cells (blastomeres) and the cell behaviors that are displayed by them: cell division, cell polarization and cell adhesion. Besides that, we added two cell behaviors not present in the original framework but that are supposed to relevant for spiral cleavage (Sachs´ rule and cortical rotation). All the developmental mechanisms and biomechanical properties that are present in the EmbryoMaker framework but are not displayed by the blastomeres (or are not relevant in the context of early cleavage) were ignored. We call SpiralMaker to this modification of EmbryoMaker (everything we describe about the EmbryoMaker is included also within the SpiralMaker and then we describe features specific of the SpiralMaker).

Each cell is made of subcellular elements (nodes), which can be conceptualized as spherical portions of a cell occupying a physical volume. The number, size and position (in a continuous 3D space) of nodes changes according to the model dynamics. As a result cells move and change their size and form. Changes in the position and shape of cells configure the overall changes in an embryo's morphology. Since in most spiralians cleavage takes place without cell growth, the number of nodes in the SpiralMaker is kept constant during each simulation. The number of nodes in a cell and the number of cells varies due to cell division. Motion is computed in continuous time and 3D space by solving a system of differential equations assuming Langevin over-damped

dynamics and some amount of noise.

All calculations, including molecule concentrations and diffusion, are made exclusively in the nodes. The spatial distribution of those nodes represents the embryo's morphology and within each cell it represents cell shape. Each node has a set of mechanical properties and can contain different types of molecules. These properties are numerical values that affect the forces acting between nodes and are affected by the molecules present in the nodes. Cells also have properties whose values are affected by the molecules present in each of their nodes.

Cell behaviors are implemented in the model as specific rules of manipulation of these nodes and their distribution between cells. The EmbryoMaker software implementing the model is available for download (http://www.biocenter.helsinki.fi/salazar/software.html). SpiralMaker is available from the same site.

2 List of model elements:

2.1 Subcellular elements and nodes.

Subcellular elements are the smallest functional entities implemented in the model. They represent a physical portion of a cell. Mesenchymal subcellular elements are implemented as spherical elastic volumes, as has been done in the Subcellular Elements Model or SEM (Newman, 2005; Sandersius amd Newman, 2008), and from now on we will call them nodes.

Node properties are named by a lower case *p* and a three letter superindex specific of each property and a second subindex for the specific node (e.g: p_i^{EQD} is the property “equilibirum distance” for node *i*; notice the superindex is not p_i multiplied *EQD* times).

2.2 Cells.

Cells are functional entities in the model that can change their shape and perform a wide range of cell behaviors (see Section 5). Cell properties are named by a upper case *P* and a three letter superindex specific of each property and a second subindex for the respective cell. Thus, for example, P_i^{PHA} is property “PHA” of cell *i*.

Each cell is composed of one or several subcellular elements and the shape of the cell, thus, is given by the relative positions of those elements. The number of nodes in a cell in a given moment depends on how many nodes it had initially and on whether it has divided or not and on two cell properties. The fact that cells are made of nodes permits cells to have internal spatial

asymmetries, that is, different nodes in a cell can have different properties (either different mechanical properties or different amounts of different regulatory molecules).

2.3 Regulatory molecules.

The EmbryoMaker can simulate any gene network with any number of models but in the SpiralMaker we only consider three regulatory molecules and these do not interact with each other:

Molecule A: It forms an Animal-vegetal gradient, in which its concentration in each node is linearly proportional to its distance to the animal pole. This molecule specifies the animal-vegetal axis required for the cell polarization rule.

Molecule B: It forms an Animal-vegetal gradient, in which its concentration depends non-linearly (as specified by the parameter k^{asym}) on its distance to the animal pole. This molecule specifies how asymmetric cell division is.

Molecule C: It is expressed only in those nodes that are or become in contact with nodes of other cells. This molecule is required for the cell-cell contact rule. Since nodes change in position this molecule can become expressed or repressed as nodes dynamically change their neighbors over simulation time.

2.4 Global model parameters.

These are numerical values which, like the molecular parameters, do not change during a given model simulation but that can be set to different values in different simulations of the model and do not correspond directly to any distinct genetic entity. These parameters include things such as the temperature and *logic model parameters* that specify some details about how the model is actually numerically implemented. These logic model parameters are identified by an *L* with a subindex specific to each parameter. Other model parameters are represented by an *M* with a subindex specific to each parameter (see section 8).

2.5 Initial conditions.

Those are the numerical values of all the nodes, and all the node and cell properties at time

zero of a simulation (thus including its location in 3D space and the amount of each regulatory molecule present in each node). They are simply the stage in development from which we want to start simulating development. Each initial condition is thus what we call a developmental pattern (Salazar-Ciudad et al., 2003) and the model dynamics transform the initial pattern into a different later pattern according to the model parameters.

Since the first two divisions are very similar among many phyla (spiralian and non-spiralian (Gilbert and Raunio, 1997)), in the SpiralMaker the initial conditions are four cells (blastomeres) that form a square in the XY plane and come from a single spheric oocyte. In the first experiment (see section B), the total number of nodes in the system is 1250, and the four blastomeres are equally sized (symmetric initial conditions). In the second experiment (See section C), the total number of nodes in the system is 2500, and both symmetric and asymmetric initial conditions were considered. In the case of asymmetric initial conditions, we estimate the relative volume of the four unequal macromeres after measuring their relative radii in real embryos as found in a bibliography search (Wilson, 1892; Newell, 1948; Freeman and Lundelius, 1982). In the case of *Arenicola* and *Nereis,* we estimate that the D-blastomere is about four-fold larger than the A-C blastomeres, so the number of nodes in the four initial blastomeres were adjusted to this proportion (see Fig. 1). In all initial conditions, each node expresses:

1) Animal-vegetal gradient 1: Molecule A with a concentration that is linearly proportional to how close a node is to the animal pole. This molecule specifies the animal-vegetal axis required for the cell polarization rule.

2) Animal-vegetal gradient 2: Molecule B with a concentration that depends non-linearly (as specified by the parameter k^{asym}) on how close this node is to the animal pole. This molecule specifies how asymmetric is cell division.

3) Cell-cell contact gradient: Molecule C that is expressed only in those nodes which are or get in contact with nodes of other cells. This molecule is required for the cell-cell contact rule. Since nodes change in position, we force this molecule to be expressed or repressed as nodes dynamically change their neighbors over simulation time.

Many real spiralian embryos are surrounded by a hyaline layer, which may exert a compressive effect on blastomeres. This outer cover is not strictly required for the spiral pattern (micromeres of real spiralians do twist in a coherent manner even if the hyaline layer has been

removed (Kuroda et al., 2009)). Moreover, that wouldn't be able to explain that, for a given species, all embryos always show either clockwise or counterclockwise twisting (because the compressive effect, as cell-cell adhesion, is not chiral).

However, by limiting the available physical space for the micromeres, this outer cover could facilitate the correct positioning of the micromeres in the furrows between macromeres (Kajita et al., 2003). Thus, we also ran some simulations with an external coverage or eggshell surrounding the blastula. This eggshell was spherical, and its center was the origin of coordinates (x=0, y=0, z=0).

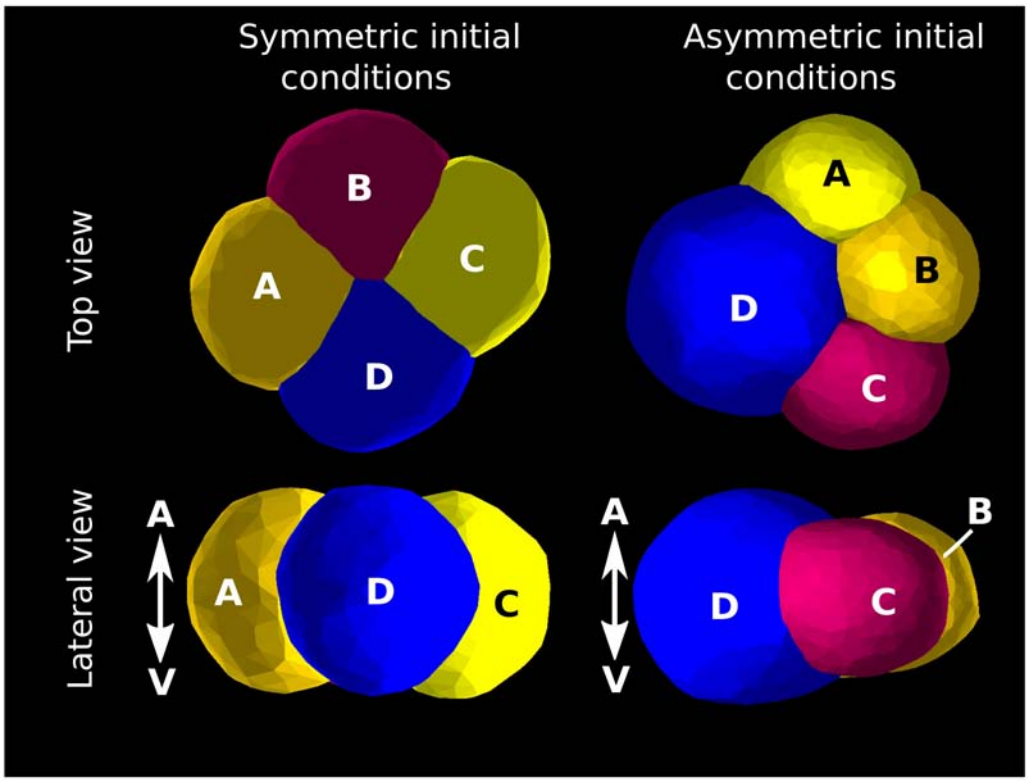


Figure 1: Initial conditions. The picture shows the two initial conditions (symmetric and asymmetric) that have been used in the model. In the symmetric initial conditions (left), the four macromeres (A-B) are equally sized, whereas in the asymmetric initial conditions the D blastomere is four-fold larger than the others, as in the case of some real Spiralians (Arenicola). In both cases, the centroids of each cell lies approximately in the same XY plane, and a cross furrow (a small point of contact between opposing macromeres) appears as a consequence of cell adhesion. Cell surfaces are drawn by means of a smoothed triangular mesh that links the external nodes of each cell. All our simulations start after the second blastomere division because before this stage, the spiral cleavage proceeds as in other, non-spiral cleavage types.

3. Mechanical Forces

3.1 Node neighboring and basic biomechanical interactions:

Nodes have a size, specified by a radius in the spheres. Two nodes adhere to each other if they come into contact (Fig. 2A), that is, if their distance is smaller than the sum of their radii. The radius is a node property, p^{ADD}. Adhesion brings these nodes closer, further decreasing their distance until the equilibrium distance between nodes is reached. This represents the generic property of cell adhesion. If the nodes are from different cells this adhesion can be increased or decreased by adhesion molecules expressed at the nodes, as we will later explain.

Cell parts such as nodes represent physical objects and thus two nodes cannot occupy the same spatial location. Nodes, thus, have a second radius, node property p^{EQD}, and two nodes start repelling each other if the distance between them is shorter than the sum of their p^{EQD} (Fig. 2A). If the distance between two nodes is exactly equal to the sum of their p^{EQD} radii then an equilibrium is reached in which these nodes neither attract nor repel each other.

From these forces follows that the movement of node *i* due to its interaction with other nodes is:

$$\frac{\partial \vec{r}_i}{\partial t} = \sum_{j=1}^{j=n_d} f_{Aij}\ \hat{u}_{ij} \quad (1)$$

Where n_d is the number of nodes in the embryo, $\vec{r}_i$ is the position in three-dimensional space of node *i*, *t* is time (the model uses continuous time), f_{Aij} is the force modulus and $\hat{u}_{ij}$ is the unit vector between node *i* and node *j* for spherical nodes and an analogous property for cylinders. We assume that most developmental processes take place within highly viscous media (Purcell, 1997), thus we calculate movement through an over-damped equation of motion. The force modulus between node *i* and *j* is:

$$\left\{ \begin{array}{c} f_{Aij} = k_{ij}^{REP}\left(d_{ij} - d_{ij}^{EQD}\right) \ \ if \ \ d_{ij} < d_{ij}^{EQD} \\ f_{Aij} = k_{ij}^{YOU}\left(d_{ij} - d_{ij}^{EQD}\right) \ \ if \ \ d_{ij} \le d_{ij}^{EQD} \le d_{ij}^{ADD} \\ f_{Aij} = 0 \ \ if \ \ d_{ij} > d^{ADD} \end{array} \right\} \quad (2)$$

Where d_{ij} is the the distance between node *i* and *j*, d_{ij}^{EQD} is the equilibrium distance between node *i* and node *j* (simply the sum of the equilibrium radii, *EQD*, node properties of nodes *i* and *j*, Fig. 2A) and d_{ij}^{ADD} is the sum of the node property *ADD* of node *i* and *j*:

$$\begin{aligned} k_{ij}^{REP} &= p_i^{REP} + p_j^{REP} \\ k_{ij}^{YOU} &= p_i^{YOU} + p_j^{YOU} \\ d_{ij}^{EQD} &= p_i^{EQD} + p_j^{EQD} \\ d_{ij}^{ADD} &= p_i^{ADD} + p_j^{ADD} \end{aligned} \quad (3)$$

p^{REP} and p^{YOU} are bio-mechanical properties of the nodes that specify how strong per unit distance the repulsion and elasticity forces respectively between a given pair of nodes are. If *i* and *j* belong to different cells, k^{ADH} is used instead of k^{YOU} and k^{REC} is used instead of k^{REP}. That is, we implement intercellular adhesion as an elastic force between cells. p^{REC} is different from p^{REP} because naturally, cells may more strongly resist incoming matter from other cells than from the same cell.

$$\begin{aligned} k_{ij}^{REC} &= p_i^{REC} + p_j^{REC} \\ k_{ij}^{ADH} &= p_i^{ADH} + p_j^{ADH} \end{aligned} \quad (4)$$

If node density in space and the p_i^{ADD} of nodes are both large it becomes possible that two nodes could interact even if there are nodes between them. To avoid that unrealistic situation the EmbryoMaker allows for three alternative algorithms to determine which nodes can effectively interact. Each simulation should be run with only one of these alternative methods. In the SpiralMaker we choose the Delaunay Method, in which a tesselation of the 3D space is performed by the Delaunay triangulation algorithm (as in Delile et al., 2013), taking each node as a vertex. Then only the nodes that are connected by an edge in this tesselation and are at a distance smaller than the sum of their p^{ADD}, as above, interact.

3.2. Node movement and noise

In addition to the movement equation defined in the previous section there is some noise in node movements. At each time step, a proportion M_{NOI} (a model parameter) of the nodes are chosen

at random and are tentatively moved in a random direction for a random distance between 0 and p_i^{DMO} , a mechanical property of each node. For each node the potential mechanical energy is calculated, by integrating the same force equations shown in section 3.1, in the new position. If the potential energy in the new position is smaller than in the old position the movement is accepted. If not, the movement is accepted with a probability proportional to the difference in potential energy between the new and old positions and inversely proportionally to a temperature parameter, model parameter M_{TEM}, plus a node property defining the node's propensity to movement (p^{MOV}),

$$\left\{ \begin{array}{l} P_{accept} = e^{-\frac{U_{after} - U_{before}}{M_{TEM} + P_i^{MOV}}} \quad if \quad U_{after} - U_{before} > 0 \\ P_{accept} = 1 \quad if \quad U_{after} - U_{before} \leq 0 \end{array} \right\} \quad (5)$$

where P_{accept} is the probability of realization of the movement, U_{before} is the potential energy in the node position before movement and U_{after} is the potential energy after the movement. If the movement is not accepted the node is put back to its old position. This energy biased noise reflects the fact that noise can affect nodes' positions but it is unlikely to bring nodes into very energetically unfavorable positions (e.g. noise is very unlikely to bring a node from a cell inside another cell). This is a standard way to implement noise in many physical and biological systems (such as in SEM and in the Pott's model (Graner and Glazier, 1992).

At the level of cells and nodes this noise property, p^{MOV}, reflects in part the tendency of cells to temporarily extent and retract cytoplasmatic projections (filopodia, pseudopodia and related structures) into the extracellular space. The likelihood of a pseudopodium retracting after being extended depends on whether it finds a suitable strong binding site (either in other cells or in the substratum). Also different types of cells tend to have pseudopodia of different lengths and tend to extend them with different frequencies. In individually migrating cells, in addition, the binding of those extensions is also relatively unstable so that cells can dynamically move over space. In our model this is captured by the p^{MOV} and p^{DMO} node properties . The movement of a node by noise can be represented as this node being the tip of a pseudopodium. p^{DMO} specifies how long pseudopodia can extend before being retracted and p^{MOV} specifies how labile this node binding is (the effect is simply to add to noise in eq 5). Each node would then bind according to its p^{ADH} , plus the amount adhesion molecules expressed both in that node and in the one it is making contact with.

In the simulations in which an outer cover (or eggshell) surrounds the blastula, if a node moves out of the eggshell, it is slightly displaced towards the center of the coordinates until it is

again within the eggshell.

4. Node properties

Most node properties have already been described in the section about mechanical forces. In the SpiralMaker these properties, and their values (which have been kept fixed for all simulations, except for p^{ADH}) are:

1- Intercellular adhesion: $p^{ADH} = [10.0, 24.0]$

2- Intracellular elasticity: $p^{YOU} = 30.0$

3- Cell compressibility to nodes from the same cell: $p^{REP} = 500.0$

4- Cell compressibility to nodes from a different cell: $p^{REC} = 500.0$

5- Filopodia extensibility: $p^{DMO} = 0.001$

6- Filopodia unstability: $p^{MOV} = 0.01$

7- Equilibrium radius: $p^{EQD} = 0.25$

8- Adhesion radius: $p^{ADD} = 0.475$

5. Cell behaviors

5.1 Cell shape change.

Cell morphology is determined by the size and relative position of the nodes in a cell. Thus, cell morphology can change due to passive processes, such as deformation by mechanical stresses. That way, cell shape change results from the model biomechanics itself and, contrary to other cell behaviors, is not directly controlled by a single (and tunable) parameter of the model.

5.2 Cell polarization and internal cell asymmetries:

The direction of cell polarization can be affected by clues in the cell's surroundings. This can happen either by short-range (Lu and Johnston, 2013) or long-range signals (Green et al., 2008)) or by inheriting an asymmetric cytosol, for example from the zygote, with an intra-cellular molecular gradient (Day and Lawrence, 2000; Lu and Johnston, 2013). In the case of molecular gradients the direction of polarization would be determined by the direction of the gradient (this is the direction

in which the concentration of the factor relevant for polarization decreases the most) (Day and Lawrence, 2000). In our model the polarization of a cell *h* is described by a 3D vector, $\hat{P}_h^{POL}$, a cell property. This vector arises from the asymmetrical distribution of regulatory molecules between the nodes within a cell (molecule A involved in cell polarization rule or molecule C involved in cell-cell contact rule, see Section 2.3).

In the first case (polarization rule) we assume there is a molecular gradient arising from the animal pole in the zygote (Raven, 1967; Gilbert and Raunio, 1997), and that this gradient is inherited without modification by each successive daughter cell. In the second case (cell-cell contact rule) the polarizing molecule C is expressed only in those nodes which are or get in contact with nodes of other cells. Since nodes change in position, we force this molecule to be expressed or repressed as nodes dynamically change their neighbors over simulation time.

First, for each cell *h*, a specific polarization vector is calculated for each molecule ($\hat{P}_h^{POLA}$ for molecule A and $\hat{P}_h^{POLC}$ for molecule C):

$$\begin{aligned} \vec{P}_h^{POLA} &= \sum_{i=1}^{n_h} \left(g_{iA} - g_{cA}\right)\vec{r}_i \\ \vec{P}_h^{POLC} &= \sum_{i=1}^{n_h} \left(g_{iC} - g_{cC}\right)\vec{r}_i \end{aligned} \quad (6)$$

Where g_{iA} is the concentration of molecule A in the node *i*, g_{cA} is the concentration of molecule A in the node closest to the centroid of cell *h*, n_h is the number of nodes in cell *h* and $\vec{r}_i$ is the position vector of node *i* (the same for molecule C). This is simply an average of each node position weighted by its concentration of each molecule, compared to that of the most central node. This vector is then divided by its module to find the polarization vector itself (the unit vectors $\hat{P}_h^{POLA}$ and $\hat{P}_h^{POLC}$).

After that, we calculate a single polarization vector of the cell *h*, $\hat{P}_h^{POL}$, as a weighted average of the polarization vectors for the molecules A and C:

$$\vec{P}_h^{\mathrm{POL}} = w_{POLA}\hat{P}_h^{POLA} + w_{POLC}\hat{P}_h^{POLC} \quad (7)$$

Where w_{POLA} and w_{POLC} are parameters of the model, whose values (in experiment 1) were 0,0.5 and 1. w_{POLA} and w_{POLC} express, respectively, the relative contribution of the polarization rule

and the cell-cell contact rule in order to determine the polarity of the cell. As before, this vector is divided by its module to find the polarization unitary vector itself $\hat{P}_h^{POL}$.

This polarization vector, that arises from asymmetries in the distribution of some molecules within the cell, will be used to determine the direction of cell division (see next section).

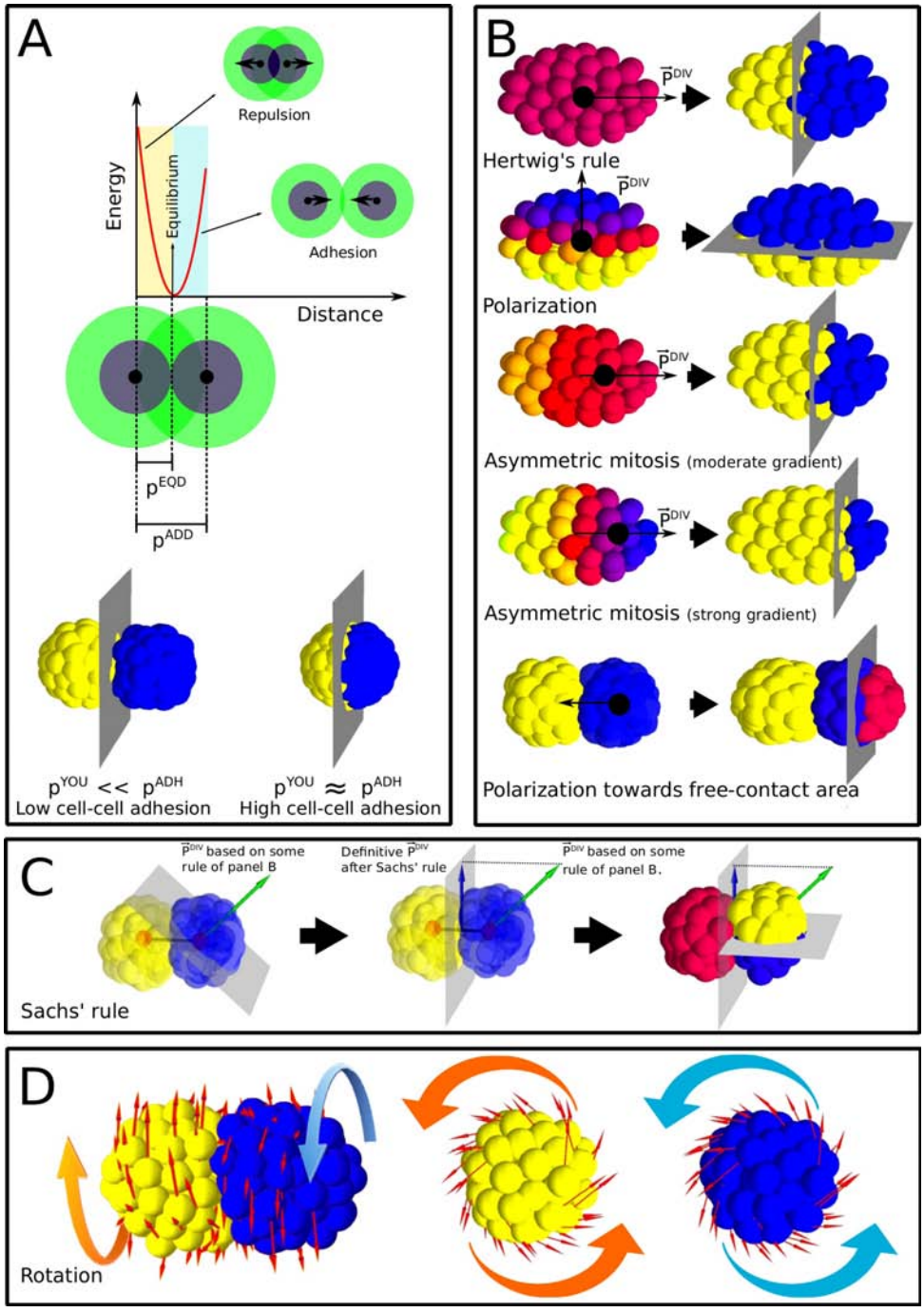


Figure 2 (From the previous page): Basic depiction of the model. A) Up: Mechanical interaction between nodes are determined by the distance between their centers and their equilibrium distance (p^{EQD}), as long as they were closer than their radius of interaction (p^{ADD}). Bottom: Effect of varying cell-cell adhesion (p^{ADH}). B) Different biological rules that cells use to specify the direction of cell division and have been implemented in the model. These rules can be combined. Left panel shows a hypothetical flattened blastomere just before mitosis (except in the polarization towards free-contact area rule, where two blastomeres appear). Different node colours represent gene concentration: high gene concentration is denoted by purple. Right panel shows the outcome of cell division under each rule (see text for a detailed explanation of the rules). Division plane (grey) separates the nodes belonging to each daughter cell (blue and yellow nodes). The division plane is perpendicular to the direction of cell division P^{DIV} (slim black arrows) and contains the division point (black dot). C) Sachs' rule. Direction of cell division takes place at right angles to the previous cell division. The prospective direction of cell division (green arrow) is projected in the previous division plane (central figure). This new vector (blue arrow) determines the actual direction of cell division (right). D) Blastomere rotation. Just after each cell division, blastomeres rotate around the axis connecting sister cells. Although the spin of rotation seems to be contrary to a external observer (left), each blastomere (center and right) do actually rotate the same direction (counterclockwise in this case) relative to each other. Small red arrows show the direction of displacement of each node during rotation.

5.3 Cell division.

Cell division proceeds by splitting the cell's nodes into two daughter cells along a cell division plane. Nodes in one side of the plane become part of one daughter cell and cells in the other side become part of the other daughter cell. The cell division plane of cell *h* is defined by the direction of the division and a division point. The direction of division is described by a vector we call $\vec{P}_h^{\mathrm{DIV}}$ that is perpendicular to the division plane. The division point is the place within the cell body where the division plane is placed (see Fig. 2B). Note that, geometrically, a point and a vector unequivocally define a plane.

5.3.1 Symmetric cell division:

Cell division is implemented by splitting an existing cell into two new daughter cells. In symmetric division both daughter cells inherit roughly the same number of nodes. In the model, as in the early cleavage of many spiralians (Cather, 1963), cell division occurs simultaneously in all blastomeres. In our model, the time between division events encompasses about 500 iterations. This time lapse is enough for each cell to reach its equilibrium shape and position within the blastula.

In nature, it is often the case that the plane of division is normal to the longest axis of the cell, what is commonly referred as Hertwig's rule (Minc et al., 2011; Minc and Piel, 2012). It has been suggested that this occurs because tension in astral microtubules depends on the contact angle between the microtubule tips and the cell's surface. By simple geometry it occurs that this angle is smallest at the cell apex (the cell apex is the cell boundary along the longest axis of the cell). This results in an attachment of the astral microtubules at the apex which exerts a stronger tension than the microtubules attached elsewhere, thus leading to the alignment of the mitotic spindle along the longest axis of the cell (Théry and Bornens, 2006; Gillies and Cabernand, 2011; Minc and Piel, 2012). The longest cell axis (Hertwig vector: vector $\vec{P}_h^{HER}$) is calculated by means of a 3D linear regression of nodes' positions.

On the other hand, in many developmental systems the plane of cell division is normal to the polarization axis of the cell. This polarization axis is, in turn, defined by the direction of a molecular gradient across the nodes in a cell (as specified in section 5.2).

In this model, the actual division vector (the vector normal to the plane of division) is calculated as a weighted average of the Hertwig and polarization vectors:

$$\vec{P}_h^{\mathrm{DIV}} = \left(1-w_h\right)\hat{P}_h^{POL} + w_h \hat{P}_h^{HER} \quad (8)$$

The weighting factor for any cell h, w_h, is a parameter of the model, specifying the relative wight of Hertwig's rule in determining the direction of cell division. Thus, if w_h =1 the division vector is equal to the Hertwig vector and if W_h=0 it is equal to the polarization vector. w_h=0.5 was also considered (for experiment 1).

The actual plane of division is normal to the division vector $\vec{P}_h^{\mathrm{DIV}}$ and passes through the centroid of the cell. This plane splits the cell into two and nodes in one side of the plane are assigned to one daughter cell and nodes in the other to the other.

5.3.2 Asymmetric cell division.

In asymmetric division the size of the two daughter cells is different (one daughter cell has more nodes than the other). In this case, the position of the division plane along the division vector, $\vec{P}_h^{\mathrm{DIV}}$, does not pass by the physical center of the cell (the centroid), but depends on the spatial distribution of certain regulatory molecules (in SpiralMaker by molecule B) within the cell. The

division plane is placed at the point in the axis defined by $\vec{P}_h^{DIV}$ where the sums of the concentrations of molecule B of nodes at each side of the plane are equal. Thus, the more skewed the distribution of those molecules, the more asymmetric is the cell division. If the gene product distribution is uniform then the plane of division appears in the centroid of the cell.

If the division is very asymmetric and the cell has not a very regular shape then daughter cells with isolated nodes can be produced (that is nodes in a cell not having physical contact with each other). Since this situation is a biologically unrealistic outcome of cell division, the physical integrity of potential daughter cells is checked before cell division. If a daughter cell has unconnected nodes, the division plane is moved again to a position closer to the centroid, until all the nodes of the two daughter cells have been connected.

In the SpiralMaker, the regulatory molecule (B) affecting asymmetric cell division forms a gradient along the animal-vegetal axis, since early blastomeres are known to often have such A-V gradients (Gilbert and Raunio, 1997; Raven, 1967). The steepness of this gradient, and ultimately the relative size between macromeres and micromeres, is specified by a parameter k^{asym} . The concentration in the node *i* of the regulatory molecule is $(r_{iz})^{kasym}$, being r_{iz} the position of the node *i* in the z coordinate (this position is always larger than zero).

The steeper the gradient the more different is the size of daughter cells. If $k^{asym}>1$, cell division becomes more asymmetric towards the animal pole, whereas when $k^{asym}<1$ blastomeres towards the animal pole are more equally-sized than those located in the vegetal pole due to the saturating profile of the gene expression pattern along the AV axis. Ten values of k^{asym} have been considered; five smaller than one: (0,1/5,1/4,1/3,1/2), and five greater: (1,2,3,4,5).

5.4 Cell adhesion:

Cell adhesion is integrated in the mechanical part the model (see Section 3). Each node has a basal adhesivity plus the one given by the expression of adhesion molecules, which depends on the affinity of the adhesion molecules expressed in each node. Adhesion is an intrinsic feature of the EmbryoMaker and it will occur between any two nodes *i* and *j* that are closer than the distance $p_i^{ADD}+p_j^{ADD}$. In the SpiralMaker all nodes have the same radius of adhesion, p_i^{ADD} (in EmbryoMaker in general this may change due to different levels of expression of different adhesion molecules in different nodes). The strength of the cell adhesion rule (one of the "non-directional processes" used in the SpiralMaker) is specified by the parameter p^{adh} (and it is the same for all nodes). The values of this parameter are specified in Section 4.

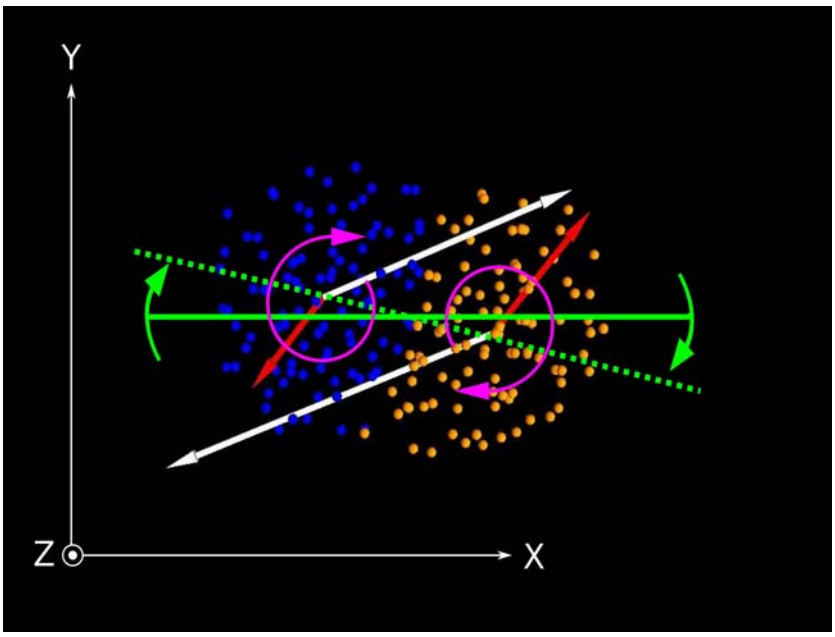


Figure 3: Cortical rotation produces a net coherent displacement of blastomeres in the same clockwise or counterclockwise direction. Simplified system of two rotating cells (viewed from the animal (+Z) pole). The blue nodes belong to one cell and yellow ones to the other cell (nodes are plotted with reduced radius for the sake of visualization). Both cells are rotating clockwise (purple arrows). The rotation axis of each cell is aligned with the animal-vegetal (Z) axis of the cell, and passes through cell's centroid. After 100 iterations, rotation in individual cells produces a net displacement of both the cells' centroids (green solid and green dashed lines represent, respectively, the lines that connect both the cells' centroids before and after rotation). This is partially explained because, although rotation is equally applied to all nodes, the nodes that contact the other cell move less than nodes not in contact with nodes of other cells. Red arrows are the average displacement vectors for nodes that contact the other cell, and white arrows for nodes that do not contact the other cell (both arrows depart from the cell's centroids). As a result, the effective rotation of the cell surface introduced by the rotation process is stronger in those areas that are not in adhesion with other cells, and the whole system rotates clockwise.

5.5 Cortical rotation

This one and the Sachs' rule are the only cell behaviors that are specific of the SpiralMaker (they are not present in the EmbryoMaker). Each cell has a rotation axis that is defined when it arises from its mother cell. This axis is the vector going from the a cell's centroid to the centroid of its sister cell. In each iteration in the EmbryoMaker some proportion of the nodes in an embryo suffers some noise. The nodes suffering this noise are randomly chosen in each iteration. In the SpiralMaker, during some iterations after cell division, a random subset of the nodes in the cell surface, the same surface nodes to which noise is applied to in a generation, suffer a small rotational

displacement. In each node this displacement is perpendicular to its cell's rotation axis and tangential to the cell surface (see Figs. 3 and 4). Geometrically there are always two vectors that are perpendicular to the rotation axis and tangential to the surface (these two alternatives represent the spin (clockwise or counterclockwise) of cortical rotation). We take arbitrarily the one in a clockwise sense. Notice that in the two daughter cells the rotation vectors are in opposite directions and that as a result both cells rotate in the same sense when seen from the center of one of the cells. When a node is performing cortical rotation, this rotation vector is summed up to the random displacement vector of the node. The resulting vector is then re-scaled to have the module of the original random displacement vector, and the node gets displaced in this direction.

Cortical rotation is parametrized according to the number of iterations after cell division is applied. Empirical evidence shows that rotation does not last more than 1/5 of the time span between consecutive mitoses (Meshcheryakov and Beloussov, 1975). Since this time encompasses in the model about 500 iterations, maximal rotation takes 100 iterations (Rot=1). Additional intermediate Rot values and no rotation (Rot=0) were also considered.

The molecular mechanics of the cortical rotation remain unclear but F-actin seems to be a major player (Danilchik et al., 2006). F-actin molecules have chiral structure and this may produce a preferential twist in the traction forces exerted by the cytoskeleton on the cortex (Henley, 2012; Tee et al., 2015). This has been proposed to produce a directional rotational flow of actin in the cell cortex (Danilchik et al., 2006; Tee et al., 2015).

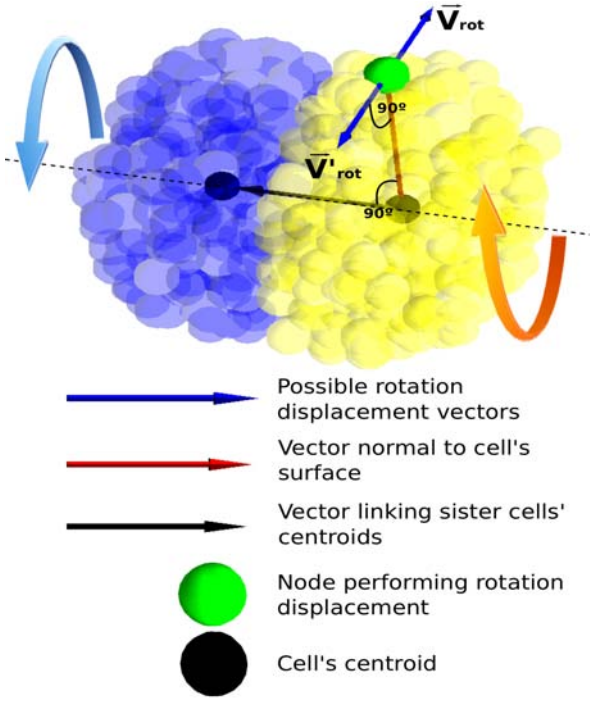


Figure 4: Cortical rotation. The picture shows two sister cells (one yellow and the other blue) performing cortical rotation after cell division. The rotation is implemented as small rotational displacements applied to a subset of nodes in the cell's surface (the same surface nodes to which noise is applied to). For the sake of clarity, we only focus on the rotational displacement (blue arrows) acting on the solid green node, belonging to the yellow cell. The black arrow is the vector linking the centroid of each cell with the centroid of its sister cell (notice that the sense of these vectors are opposite for each sister cell), and defines the rotation axis (dashed line). Red arrow is the vector linking the cell's centroid with the external node, thus being normal to the cell's surface, and perpendicular (90º) to the rotation axis. The rotational displacement applied to the node is perpendicular to the rotation axis (dashed line) and tangential to the cell's surface (that is, perpendicular to the vector normal to cell's surface (red arrow)). Two possible rotational displacement vectors (blue arrows, $\bar{\bar{V}}$rot and $\bar{\bar{V}}$' rot) satisfy these criteria. These two alternatives represent the spin (clockwise or counterclockwise) of cortical rotation. We always took, in all cells, the $\bar{\bar{V}}$rot (that produces clockwise rotation, orange arrow)

5.6 Sachs' rule:

With this rule the direction of each cell division is (or tends to be) perpendicular to the direction of the previous division (that is to the one that gave rise to the mother cell). We define the direction of previous cell division as the vector that links the centroids of the sister cells. This alone can not unequivocally determine a $\vec{P}_h^{DIV}$ vector since in 3D space there is an infinite number of

vectors that are perpendicular to any other given vector (in the this case the $\vec{P}_h^{\mathrm{DIV}}$ of the previous division). To solve that problem we first determine a $\vec{P}_h^{\mathrm{DIV}_{PRE}}$ vector based one of the rules above (Hertwig, polarization or cell-cell contact), apply it to the centroid of the cell and then project this vector into the plane in which all the vectors perpendicular to the $\vec{P}_h^{\mathrm{DIV}_{PRE}}$ of the previous division are contained (Fig. 2C). The vector resulting of this projection is the $\vec{P}_h^{\mathrm{DIV}_{SACHS}}$ of the new division.

In that sense Sachs' rule can only be applied in combination with some other rules. We also consider the situation in which Sachs' rule does not totally constrain the cell division to be perpendicular to the previous one, but just biases the cell division towards perpendicularity. In these cases, the definitive direction of cell division is defined as follows:

$$\vec{P}_h^{\mathrm{DIV}} = \left(1 - w_{SACHS}\right)\vec{P}_h^{\mathrm{DIV}_{PRE}} + w_{SACHS}\vec{P}_h^{\mathrm{DIV}_{SACHS}} \quad (9)$$

Being $\vec{P}_h^{\mathrm{DIV}_{PRE}}$ the division vector (coming from other rule(s)) before the application of Sachs' rule, and $\vec{P}_h^{\mathrm{DIV}_{SACHS}}$ the division vector modified by Sachs' rule, perpendicular to the previous cell division, as explained above. w_{SACHS} ($0<_{wSACHS}<1$) is the parameter controlling the strength of Sachs' rule, that is, the bias to the perpendicularity between cell divisions. If $w_{SACHS}=1$, Sachs' rule is totally active, and the direction of cell division will be perpendicular to the previous one. If $w_{SACHS}=0$, the Sachs' rule does not modify the direction of cell division that was previously defined by other rule(s). Intermediate situations were considered: in the first experiment $w_{SACHS}=(0,1)$, and in the second experiment $w_{SACHS}=(0,0.2,0.4,0.6,0.8,1)$.

7. Numerical integration

Differential equations are numerically integrated by the explicit Euler method. The value of δ, the integration time step, is constant over time. In our simulations δ is equal to 0.0001.

8. Model parameters:

There are some numerical model parameters that have been explained already but here we summarize them followed by the values (fixed) they have in the simulations:

-M_{TEM}: Temperature analog, this is how likely are noisy movements that are energetically unfavorable (see section 3.2). M_{TEM}=0.01.

-M_{NOI}: Proportion of nodes to which noise is applied in each iteration. M_{NOI}=0.1.

-Implementation model parameter: This is a parameter controlling the numerical implementation of the model, this is the accuracy of the model. It has no biological meaning as such.

-M_{MNN} : Maximal number of nodes any node can interact with. If there is more than that number the program crashes. These is no optimal way to avoid that effect, since these neighbors need to be stored in a temporary matrix and there are system restrictions in the size of those. In addition, there is no way to predict how many nodes a node will interact with since this is a result of model dynamics. If this value is large the program would run slower. M_{MNN} =1000.

-In the simulations in which an spheric outer cover (or eggshell) surrounds the blastula, the total volume of the cover was equal to the sum of the volumes of all the nodes (assuming that each node *i* is a small sphere with a radius equal to its equilibrium distance p_i^{EQD}) plus the volume of the empty spaces between nodes (assuming that they are randomly packed within the blastula). By geometrical considerations, it is know that randomly packed spheres fill around 64% of the total volume in which they are enclosed (Song et al., 2008). Thus, the radius $R_{eggshell}$ of the eggshell was:

$$R_{eggshell} = p^{EQD} \sqrt[3]{\frac{N_{nodes}}{0.65}} \approx 4 \quad (10)$$

By choosing this radius, the compressive effect of the outer cover is moderate (nodes can keep their equilibrium distance p^{EQD} with the neigboring nodes), but the blastula acquieres an overall rounded shape.

9. Implementation of the model in SpiralMaker.

9.1 Structure of the code. The source code is written in fortran90 and is organized in different functional fortran modules. The most relevant modules are listed below:

- general.mod.f90: declarations of the main variables that are used in common by the rest of the

modules. These are global model parameters and node and cell properties. The set of all node properties are declared in a derived type fortran 90 variable. The same occurs for cell properties. The main variables used by the other modules are a matrix of nodes and cells properties (one element per cell and node). Essentially the rest of the code is mostly operations on those matrices (including re-dimensioning them).

- model.mod.f90: manages the temporal progression of the developmental simulations and calls the subroutines in the neighboring, bio-mechanical, genetic and nexus modules (once per iteration with Euler and several times with Runge-Kutta).

- neighboring.mod.f90: contains the subroutines to calculate the neighbor relations between nodes.

- biomechanic.mod.f90: contains the subroutines that calculate the mechanic interactions and displacement of nodes.

- energy.mod.f90: contains the subroutines that calculate energy potentials for nodes that are used in energy-biased
random movements.

- genetic.mod.f90: declares the regulatory molecules and their parameters used in a specific instance of the model. It also contains the subroutines for transcription and non-transcriptional regulation.

- nexus.mod.f03: contains the subroutines that implement the molecular regulation of node and cell properties and the calls to the cell behaviors. It also contains subroutines for some simple cell behaviors.

- mitosis.mod.f90: contains the subroutines that implement cell division.

- pinta.mod.f90: This file contains two modules: A view_modifier module to control how the embryo is seen (rotation, zooming, sectioning, etc...) and a function_plotter module that contains all the subroutines that draw nodes and controls the menu. This latter module is the one including the OpenGl and glut calls.

- editor.mod.f90: contains the code required to manually edit the embryo.

- ic.mod.f90: contains a set of subroutine for simple initial conditions.

- initial.mod.f90: contains several initialization subroutines.

- io.mod.f90: contains the hard-disc input/output subroutines.

-OpenGl_gl.f90, OpenGl_glu.f90 and OpenGl_glut.f90 define fortran interfaces for the OpenGl, GLU and glut functions, and have been taken from the f03gl project (http://www-stone.ch.cam.ac.uk/pub/f03gl/index.xhtml)

9.2 Input/Output format. EmbryoMaker use a custom I/O format. The same file written by the program as output can be read as input file as well. It is basically a text file listing all the model parameters and variables that are used by the software, including node positions and gene expression levels. The names of the parameters and variables are indicated in the file, so it is possible to edit the file manually. In that sense, the editor tools of EmbryoMaker can be used to edit Input/Output files with a more intuitive graphic interface. By default, as explained in the manual, EmbryoMaker writes all the output files from a given run into a folder with a number (a different number for each run) within a folder called output.

B. In silico experiment 1: General theoretical developmental morphospace:

The application of EmbryoMaker to spiral cleavage requires: 1) Choosing as initial conditions the 4-cell stage in which cells have an intracellular molecular gradient emanating from the animal pole (presumably inherited from the zygote) 2) Dividing cells synchronously at regular time intervals 3) Using different rules to determine the direction of cell division 4) Using different steepness in the intracellular gradient affecting the asymmetry of cell division 5) Using different degrees of cortical rotation 6) Using different values of adhesion between nodes.

Except for Sachs' rule, the rotation rule and the cell-cell contact rule, the application of the model to the spiralian cleavage involves only changes in parameters that already exist in the EmbryoMaker (Marin-Riera et al., 2015).

In the model, as in the early cleavage of many spiralians (Cather, 1963), cell division occurs simultaneously in all blastomeres. Mechanical interactions between cell parts or nodes, such as the formation of adhesion points and relative cell movement, occur faster than whole-cell processes

such as cell division. In the model this is ensured by choosing long time intervals between cell division rounds. The orientation of the cell division plane is defined in several ways in the model. Each way represents one of the developmental rules determining the direction of cell division, as explained in the introduction. As explained in the introduction, these rules are:

a) Rules of division plane specification: Hertwig, Polarization, Cell-cell contact and Sachs' rules.

b) Non-directional processes: Cortical rotation, cell-adhesion and asymmetric cell division.

(See section 5 for a description of how these rules are implemented in the model).

In an initial experiment, we explored whether the combination of these rules and processes can generate the spiral pattern or not. Each simulation we ran used three different rules (plus adhesion). For this initial exploration cell division was kept approximately symmetric ($k^{asym}=1$), and the number of nodes in the blastula were 1250.

We do not consider a totally inactive ($p^{adh}=0$) cell adhesion (because it is biologically unrealistic). Instead, we ran, for each of the combinations of the remaining rules, simulations involving three different (nonzero) adhesion values. Thus, for the remaining five rules and without repetition there are $\binom{5}{3} = 5!/(3!(5-3)!)=10$ different combinations of three rules. To explore the parameter space, three values (the minimum, the maximum and an intermediate value) for each of these rules or process parameters were considered (except for the Sachs' rule, in which only a minimum and a maximum values were considered). This is $3^3=27$ (three rules, three values per rule) simulated blastulae in the five combinations of three rules which do not include Sachs' rule, and $3^2x2=18$ simulated blastulae in the remaining five combinations that include Sachs' rule (which has only two possible values). This results in (27x5)+(18x5)=225 different simulations (each one was run for three different adhesion values, yielding 225x3=675 combinations). The combinations including cortical rotation, Sachs' rule and adhesion were discarded because they lacked a vector defining the direction of cell division.

The rules of division plane specification (Hertwig, cell polarization and cell-cell contact rule) have parameter values between 0 and 1. These values represent the relative weight of each of these rules (normalized by the number of rules involved) in determining the direction of cell division $\vec{P}_h^{DIV}$.

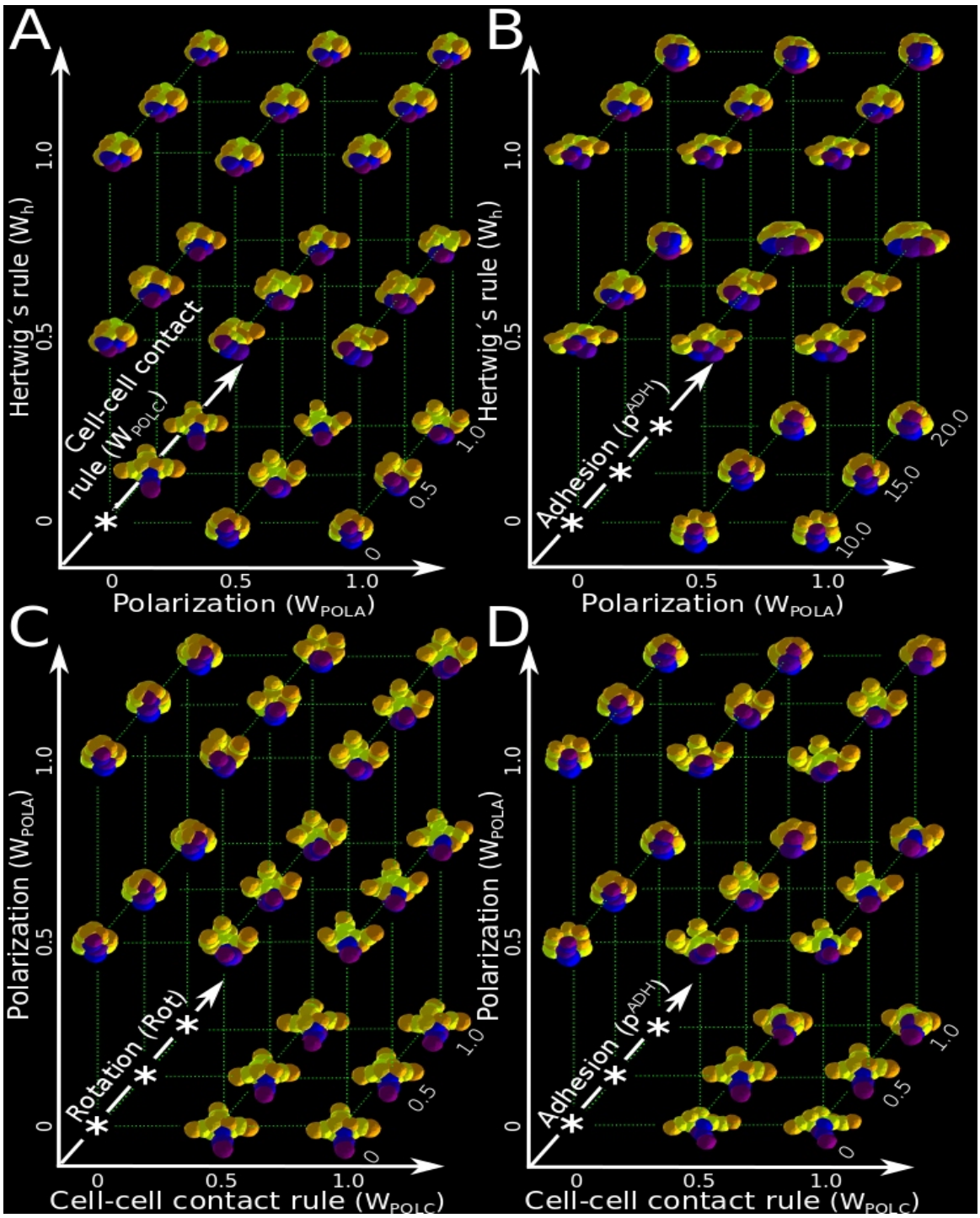


Figure 5: Example combinations of rules that do not produce spiral patterns. The resulting blastulae lack one or more of the main features that define the spiral cleavage pattern, namely: blastomeres organized in groups of four cells (quartets) along the A-V axis, at least one pair of the opposing blastomeres closer to the animal pole (that is, $1a^1$ and $1c^1$, or $1b^1$ and $1d^1$ blastomeres) contacting each other, sister blastomeres obliquely positioned in respect to the A-V axis, and blastomeres within each quartet displaced in the same direction (either all to the right or all to the left of their sister blastomere) following the left-right alternation after the 4-cell stage. Asterisks indicate hypervolumes of the morphospace that are not definable (where the combination of rules does not unambiguously specify the direction of cell division). All blastulae are drawn from a slightly lateral animal view, with the animal pole to the top.

Other, high dimensional combinations of rules (eg. four or five rules by simulation) were not considered. We discarded this possibility because our goal (the emergence of spiral cleavage pattern) was achieved under the simpler combination of only three rules (plus adhesion).

For each combination of rules and parameters, we ran a simulation until the 16-cell stage and visually checked whether the resulting blastulae reproduced the main features of spiral cleavage pattern or not. Specifically, we assessed if the blastomeres were organized in groups of four cells (quartets) along the animal-vegetal axis, if the opposing animal micromeres contacted each other (cross furrow), if the sister blastomeres were obliquely positioned respect to the animal-vegetal axis, if the blastomeres within each quartet were displaced in the same direction (either all to the right or all to the left of their sister blastomere) and if this direction followed the left-right alternation after the 4-cell stage.

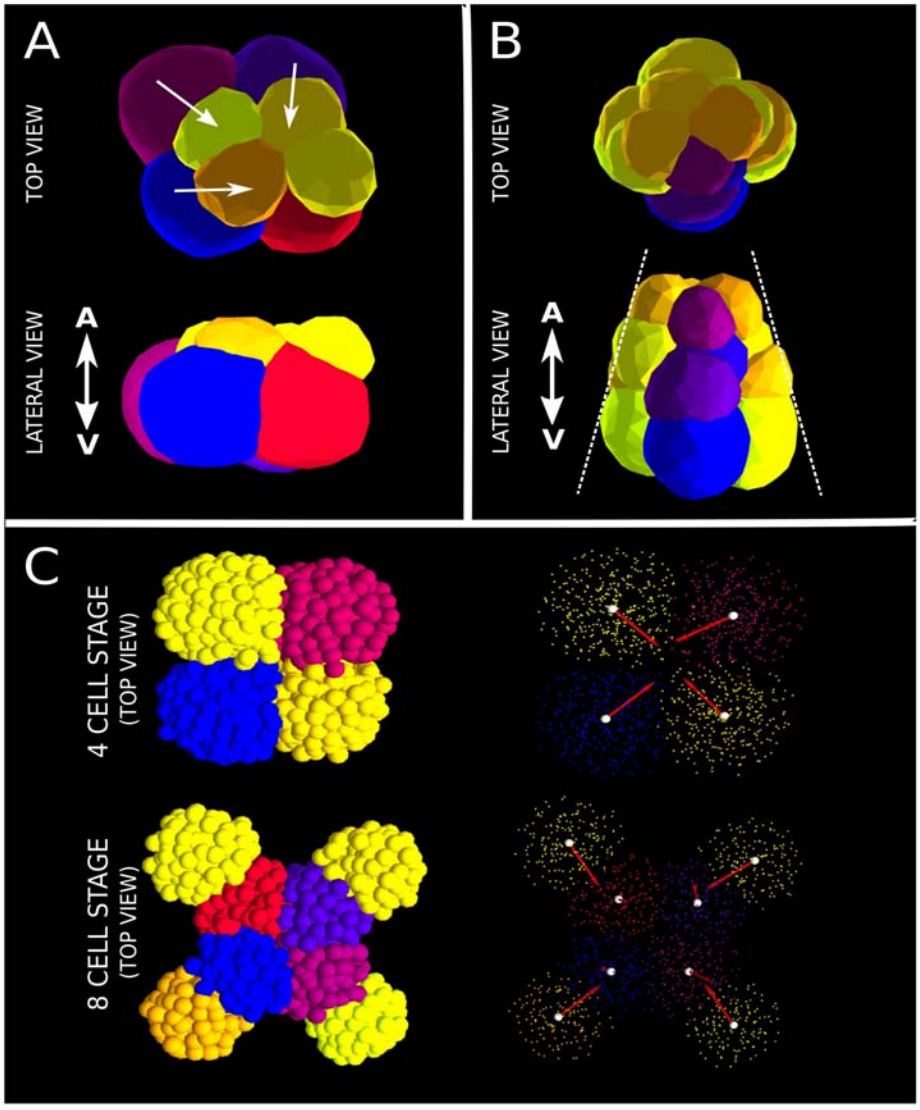


Figure 6 (From the previous page): Non-spiralian patterns arising in the model. A) Cell adhesion without cortical rotation can not produce the coordinate twist of the micromeres (all to the right, or all to the left) in the 8 cell stage. Adhesion tends to increase the contact surface between blastomeres, so each micromere tends to place itself between two macromeres. This leads to a relative twist between micro- and macromeres. However, in absence of cortical rotation, this twist is random in direction, and often the twist between two adjacent micromeres may be in a contrary direction (see white arrows). B) The adhesion and cortical rotation processes alone can not account for the spatial distribution of blastomeres in the spiral pattern after the 8 cell stage. With these rules blastomeres always divide oriented towards the animal pole and form cone shaped blastulae, with anomalous (non-spiralian) contacts between blastomeres. C) The cell-cell contact rule can not be involved in the spiralian pattern. When this rule is applied in the 4 cell-stage (upper panel), the direction of cell division (red arrows) tends to point towards the center of the blastula. In the 8 cell-stage (lower panel), in the most external cells the direction of cell division also tends to point towards the center of the blastula. That way new cells are displaced towards the periphery of the blastula The most internal (central) cells have more than one point of contact with other cells, so that their direction of cell division does not have to point to the center. In the right panels, nodes are displayed with reduced radius to visualize the cells' centroids (white spheres) and the directions of cell division (red arrows). In A-B, cell surfaces are drawn by means of a smoothed triangular mesh that links the external nodes of each cell.

C. In silico experiment 2: Morphometric comparison of virtual and real blastulae.

In a more detailed approach we took all the combinations of rules that in experiment 1 were found to produce spiral patterns and performed a more exhaustive sampling of each rule and process parameters (increments of 20% in each parameter from its minimum possible value to its maximum possible value, and 9 different adhesion values, comprising 7128 simulations, see section D). Some simulations were also run with an outer cover of radius 4 (see next section). For this experiment, the number of nodes per simulation was 2500, and simulations were run until the 32-cell stage. In addition, asymmetric cell division was included (see below and SI section 5.3.2), as well as intermediate strengths of Sachs' rule (a bias to perpendicularity between consecutive cell divisions).

In order to compare the simulated blastulae with those of specific species we followed two alternative approaches (because a comprehensive information about the 3D positions of all cells and cells´ structures within the blastula is usually not directly available from publications).

1. Relative volumes between blastomeres.

In this approach, we considered the relative volumes between the different blastomeres

along the animal-vegetal axis. We recorded these data from real embryos of three gastropods mollusks *Crepidula*, *Planorbella* and *Lottia*. For these species, high-resolution microphotographies (Zeiss confocal microscope) were obtained from fixed embryos (stained with anti-ß-tubulin E7) in the 16-cell stage (See below for a detailed description of the preparation of the embryos).

From these high-resolution microphotographies of the 16-cell stage blastulae of these species, the relative volumes of blastomeres were obtained. We did it for at least one blastomere of each quadrant (in these species, the blastomeres in the 4-cell stage are equally sized as also are the four blastomeres within each quartet). Since blastomeres are not perfectly spherical, we used the ellipsoid volume formula, which requires three orthogonal measures (d_1, d_2 and d_3) of the ellipsoid (blastomere)´s diameters. Thus, we first measured with a standard Image Processing Program (ImageJ (Schneider et al, 2012)) the three maximum visible blastomere diameters that were perpendicular (or nearly perpendicular) between them (d_1 and d_2 aligned respectively in the X and Y axes in animal view, and d_3 in the Z axis in lateral view). Due to the compact blastomere arrangement, it was often the case that some of these diameters could not be readily obtained because part of the blastomere was hidden between other blastomeres, In this case, we averaged the missing diameter from the other two. In the considered embryos, blastomeres belonging to each quartet are equally sized (each quartet is the set of four blastomeres, simultaneously released, that are in the same relative position along the animal-vegetal axis). Thus, we measured only some diameters within each quartet to get the average values ($\bar{d}_1$, $\bar{d}_2$ and $\bar{d}_3$) of each quartet. The average absolute volume of each blastomere within given quartet "*x*" in the 16-cell stage is then:

$$V_A(x)=\frac{1}{6}\pi\left(\bar{d}_{x1}\bar{d}_{x2}\bar{d}_{x3}\right) \quad (11)$$

For each blastomere in a species and its closest simulated blastula we calculated the square of the differences in relative volume. The sum of those for all the blastomeres in a blastula was used as a measure of the overall similarity or distance between each species blastula and its closest simulated blastula.

For this experiment, we also ran some simulations in which an outer cover (or eggshell) surrounds the blastula. We did it in order to improve the visual resemblance between simulated blastulae and the real ones (whose shape was almost spherical). When we did it, we kept minimal square differences in relative volumes.

2. Relative positions between adjacent blastomeres.

In a second approach, we quantified the similarity between the blastulae of real species and the bastulae in our theoretical morphological morphospace in terms of the relative positions between adjacent blastomeres. That is, by measuring which proportion of the contacts between specific blastomeres (according to their lineage) are the same between each simulated blastula and each species blastula. .

For the 16 and 32-cell stages of spiral cleavage there is a clear nomenclature and homology for blastomeres according to their cell lineage (for example the cell $1d^2$ is the cell arising from the second division of the D macromere that is closer to the animal pole) (Gilbert and Raunio, 1997). Thus, we recorded, for each simulation, which blastomeres (identified according to this cell lineage) are in physical contact with which other blastomeres. The blastomere contacts between each spiralian blastula in those stages was summarized in a table, or matrix, of 0s and 1s (1 for contact and 0 for non-contact) that has the same size and shape in all species and in all *in silico* blastula that have a spiral pattern (each row and column in this matrix is simply one of the 16 cells, e.g: $1d^2$)).

For four different spiralian species, those matrices were manually extracted from bibliographic data, specifically from studies presenting apical and lateral views of 16-cell stage embryos with clear depictions of the contacts between adjacent cells (these limits between blastomeres were not so clearly defined in our samples of real embryos, so they were discarded for this approach). Data for *Trochus* (gastropod snail) were extracted from (Robert, 1902; Goulding, 2009); for *Carinoma* (nemertean worm) from (Maslakova et al. 2004); for *Nereis* (polychaete worm) from (Wilson, 1892; Freeman and Lundelius, 1982) and for *Arenicola* (polychaete worm) from (Newell, 1948). In two of these species (*Arenicola* and *Nereis*), the D-blastomere (a mesoderm precursor) remains visibly larger than the other ones due to an asymmetrical inheritance of cytoplasmatic determinants in the 4-cell stage (Freeman and Lundelius, 1982). Thus, we also ran simulations considering this asymmetry in the size of the blastomeres as initial conditions. We estimated the relative volume of these four unequal macromeres after measuring their relative radii in real embryos as found in (Wilson, 1892; Newell, 1948; Freeman and Lundelius, 1982). In the case of these polychaetes, we estimated that D blastomere is about 4.16 times larger than the A-C blastomeres.

Then, for each species, we screened the theoretical morphological space for simulated embryos whose similarity to the real ones was >90% (measured as the proportion of blastomere contacts identical between simulated and real blastulae). This is, the sum of the differences between corresponding elements in two tables. This sum was then divided by the number of possible cell-cell contacts (16x16=256) and multiplied by 100 to get the percent similarity. We also obtain, for

each species, its relative occupation of the morphological space as the number of simulated cleavage patterns with a similarity of at least 90% divided by the total number of points simulated in the morphological space.

3. Sample collection and microphotography.

Adult *Crepidula fornicata* were collected near Woods Hole, MA, by the Marine Resources Center at the Marine Biological Laboratory. Embryos were reared as described in Henry et al. (2006). A breeding population of *Planorbella duryi* is maintained in freshwater tanks at 25 °C. Sexually mature *Lottia gigantea* were collected in Los Angeles, California, during the breeding season, and in vitro fertilizations were performed as described in Grande and Patel (2009). Embryos of the three species were fixed for one hour at room temperature in a 3.7% solution of ultrapure formaldehyde (Ted Pella, Inc., Redding, CA) dissolved in filtered sea water (freshwater for *P. duryi*). Following fixation, embryos were rinsed with three sterile 1X PBS washes (1XPBS:1.86mM NaH2PO4, 8.41mM Na2HPO4, 175mN NaCl, pH 7.4), followed by three washes in 50%, 70% and 100% methanol. Embryos were stored in 100% methanol at -80°C. For the immunohistochemistry reactions, embryos were washed stepwise from methanol into PBS to remove any trace of methanol and then washed for 1 hour in blocking solution (1X PBS with 0,1% Tween 20 and 4% BSA). Samples were then incubated overnight at 4ºC in fresh blocking solution containing the appropriate solution (1:10) of primary antibody (mouse anti-ß-tubulin E7, Developmental Studies Hybridoma Bank, NICHD and University of Iowa). Samples were then washed several times in PBS and then incubated for two hours at room temperature in fresh blocking solution containing the appropriate solution (1:200) of secondary antibody (Alexa Fluor-555, Invitrogen) and finally washed again several times in PBS. Following immunohistochemistry reactions, embryos were incubated in a solution of Hoechst 33342 (Invitrogen) diluted 1:1000 for 10 minutes, then washed several times in PBS and finally into 80% Glycerol/20% PBS. Photographs of samples were acquired using a LSM710 confocal microscope (Zeiss). The preparation of the samples and the Confocal Microscopy was performed at the Microscopy Unit at Centro de Biologia Molecular Severo Ochoa (Spain).

*4. Empirical and simulated measures for each species blastulae***:**

In order to compare real and simulated blastulae with the “relative blastomere volumes” approach, the average diameter (d), absolute volumes (V_A) and relative volumes per blastomere for each quartet for the different spiralian species were collected (Data for *Crepidula, Planorbella* and

Lottia were extracted from fixed embryos, and for *Trochus* and *Carinoma* from bibliographic data). Their values are summarized in the next table:

		Crepidula	***Planorbella***	***Lottia***
Average diameter (quartet 2Q)		130.5 µm	51 µm	26.5 µm
Average diameter (quartet 2Q)		81 µm	34 µm	19.5 µm
Average diameter (quartet 1q1)		52 µm	29.75 µm	19.5 µm
Average diameter (quartet 1q2)		44 µm	24 µm	17.5 µm
Average absolute volume (Quartet 2Q)		1163670 μm^3	69456 μm^3	77952 μm^3
Average absolute volume (quartet 2q)		278262 μm^3	20580 μm^3	31059 μm^3
Average absolute volume (quartet 1q1)		73622 μm^3	13856 μm^3	31059 μm^3
Average absolute volume (quartet 1q2)		44602 μm^3	7238 μm^3	22449 μm^3
Average relative volume (quartet 2Q)	Real	0.74	0.62	0.48
	Simulated	0.74	0.62	0.47
Average relative volume (quartet 2q)	Real	0.18	0.19	0.19
	Simulated	0.17	0.19	0.23
Average relative volume (quartet $1q^1$)	Real	0.05	0.12	0.19
	Simulated	0.06	0.12	0.16
Average relative volume (quartet $1q^2$)	Real	0.03	0.07	0.14
	Simulated	0.03	0.07	0.13

Where: Quartet "Q" includes blastomeres 2A,2B,2C and 2D. Quartet "q" includes blastomeres 2a, 2b, 2c and 2d. Quartet "$2q^1$" includes blastomeres $2a^1$, $2b^1$, $2c^1$ and $2d^1$. Quartet "$2q^2$" includes blastomeres $2a^2$, $2b^2$, $2c^2$ and $2d^2$}. In this approach, the parameter values of the simulated blastulae that best matched with the ones of the real species according to the relative volumes of their blastomeres, and the distance between them (see SI section C1), were as follows:

	Crepidula	***Planorbella***	***Lottia***
Cortical rotation (Rot)	0.2	0.6	0.8
W_{SACHS}	0.6	0.8	0.8
Cell adhesion (p^{ADH})	18.0	24.0	24.0
k^{asym}	5.0	3.0	2.0
Eggshell radius	No eggshell	4.0	4.0
Distance relative volumes (Real, Simulated)	0.017	0.037	0.046

When real and simulated blastulae were compared by the second "blastomere contact-based" approach, the parameter values that produce each species blastula were as follows:

The cleavage pattern best matching with *Trochus* (Fig. 4B in Chapter 2, 100% similarity in cell contacts between simulated and observed blastulae) arose when cortical rotation was moderate (Rot=0.4), Sachs' rule was 100% active (w_{SACHS}=1), cell adhesion p^{ADH} was at any value in the range between 14.0 and 16.0 (the medium region of the adhesion range considered in the model; see

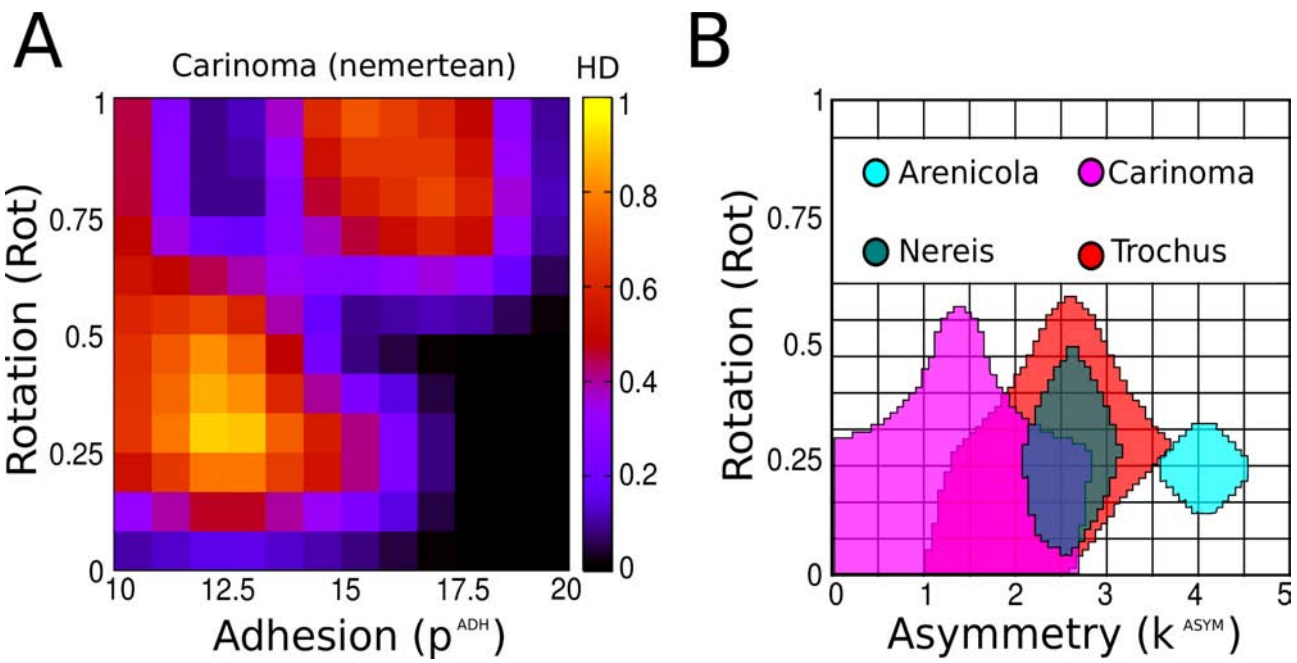
methods) and the cell division gradient had any value in the range $2.0<k^{asym}<3.0$ (moderately steep gradient, leading to “intermediate” asymmetric cell division). This species has the spiral pattern that is most common of the four species in the parameter space (see Methods) (31% of the parameter space had a 90% similarity in cell contacts with *Trochus*).

The cleavage pattern best matching the nemertean *Carinoma* (Fig. 4C in Chapter 2, 99% similarity in cell contacts) appeared for low rotation (Rot=0.2), Sachs' rule's contribution of 50% ($w_{SACHS}=0.5$), a relatively low adhesion value of $p^{ADH}=12.0$, and a cell division gradient $k^{asym}=0.5$ (this is moderate gradient in which the concentration changes only slithgly close to the animal pole and changes more abruptly close to the vegetal pole). The basic visual features of this blastula, which closely resembles the one arisen from the model, are the very rounded blastomeres (likely due to decreased cell-cell adhesion as in the model) and the existence of large and prominent micromeres (likely due to the saturating ($k^{asym}<1$) profile of the gene expression gradient along the animal-vegetal axis, that controls asymmetric cell division).

The cleavage pattern best matching with *Nereis* (Fig. 4D in Chapter 2, 98% similarity in cell contacts) was found for intermediate rotation values (Rot=0.4), Sachs' rule totally active ($w_{SACHS}=1$), adhesion in the intermediate value ($p^{ADH}=16.0$), cell division nearly symmetrical ($1.0<k^{asym}<2.0$) and initial conditions with three equal blastomeres and a blastomere 4.16 times larger than the others.

The cleavage pattern best matching with the other polychaete considered (*Arenicola*) (Fig. 4E in Chapter 2, 98% similarity in cell contacts) arises when cortical rotation has the minimal value considered (Rot=0.2), Sachs' rule's contribution is 50% ($w_{SACHS}=0.5$), adhesion has an intermediate value (is $p^{ADH}=16.0$), cell division is quite asymmetric (A-V gradient with a $k^{asym}=3.0$) and, of the four initial blastomeres, three are equal but D is 4.16 times larger than the other (as in the real embryo).

Notice that the parameter combination that leads to each blastula reminiscent of a real species is not exactly the same under the two different approaches. This is partially explained by the fact that these two approaches only capture two aspects (thought very important ones) of the general embryo morphology. Fig. 7 shows that the hypervolumes of the parameter space occupied by the different species differ in size and shape. As we can see in Fig. 2B, when5the parts of the heatmaps where the different species are found are simultaneously plotted in a 2D representation of the parameter space, we observe that some hypervolumes do overlap between each other, the amount of overlapping representing the likelihood of a particular transition between patterns. Our results show that the hypervolumes of closely related species (e.g. between *Nereis* and *Arenicola* polychaetes) do not overlap and that some species occupy a larger hyper-volume (notably *Trochus*).



Supplementary Figure 5. Spiralian parameter space A) Example heatmap of the morphospace of the simulated blastulae compared with a real blastula (Carinoma). Colors indicate in which region of the morphospace embryos similar to the real one are located. Similarity is measured as the proportion of contacts between blastomeres that are the same between real and simulated embryos. Regions where the species-specific pattern is found appear as two scattered islands in a morphospace, which shows that the structure of the morphospace is complex, suggesting an indirect relationship between developmental parameters and morphology. B) Exploration of the morphospace. Different and related species may acquire different embryo morphologies by tunning developmental parameters over evolutionary time. Colored area encompasses regions with >0.75 of similarity (smoothing was applied to the contour for the sake of visualization).

D. Parameter exploration ranges:

The strengths of the developmental rules that have been used in this work are controlled by the previously explained model parameters. We summarize the values of these parameters:

1. Hertwig's rule: It depends on the relative weight of the Hertwig vector $\vec{P}_h^{HER}$ compared with the polarization vector $\hat{P}_h^{POL}$ in determining the direction of cell division. This relative weight is controlled by the parameter W_h ($0<W_h<1$)(see section 5.3). Thus, if w_h =1 the division vector is equal to the Hertwig vector (Hertwig's rule is “totally active”) and if W_h=0 it is equal to the polarization vector. In the parameter space exploration of experiment 1 three values were considered: 0, 0.5 and

1. We only tried three values because the number of different simulations to run increases very fast with that number, as the combination D!/(N! (D-N)!), where N is the number of parameter values tried by parameter and D is the number of parameters changed in the model.

2. Cell polarization rule: This rule is parametrized according to the relative weight (w_{POLA}) of the polarization vector of molecule A $\hat{P}_h^{POLA}$ in determining the cell polarization vector $\hat{P}_h^{POL}$, which in turn will determine the direction of cell division (the molecule A forms an animal-vegetal gradient and specifies the animal-vegetal axis required for the cell polarization rule). Three values of w_{POLA} were considered for experiment 1: (0,0.5,1). In experiment 2, w_{POLA} is always equal to 1.

3. Cell-cell contact rule: This rule is parametrized according to the relative weight (w_{POLC}) of the polarization vector of molecule C $\hat{P}_h^{POLC}$ in determining the cell polarization vector $\hat{P}_h^{POL}$ (which in turn will determine the direction of cell division). This molecule C is expressed only in those nodes that are in contact with nodes of other cells, and is required for the cell-cell contact rule. Three values of w_{POLC} were considered for experiment 1: (0,0.5,1). In experiment 2.

4. Sachs' rule: The strength of Sachs' rule, that is, the bias to the perpendicularity between cell divisions is controlled by the parameter w_{SACHS} ($0<_{wSACHS}<1$). If $w_{SACHS}=1$, Sachs' rule is totally active, and the direction of cell division will be perpendicular to the previous one. If $w_{SACHS}=0$, the Sachs' rule does not modify the direction of cell division that was previously defined by other rule(s). Intermediate situations were considered: in the first experiment $w_{SACHS}=(0,1)$, and in the second experiment $w_{SACHS}=(0,0.2,0.4,0.6,0.8,1)$.

5. Cortical rotation: This rule is parametrized according to the number of iterations in which rotation is applied after cell division is applied. Empirical evidence shows that rotation does not last more than 1/5 of the time span between consecutive mitosis (11). Since this time encompasses in the model about 500 iterations, maximal rotation takes 100 iterations (Rot=1). For experiment 1, only Rot=1, no rotation (Rot=0) and intermediate rotation (Rot=0.5) were considered. In the experiment 2, Rot=(0, 0.2, 0.4, 0.6, 0.8, 1).

6. Asymmetric cell division: The relative size of daughter cells depends on the spatial distribution of molecule B within the mother cell (see section 5.3.2). Molecule B forms a gradient along AV axis, and the steeper the gradient the more different the size of daughter cells. The steepness of this

gradient is regulated by the k^{asym} parameter. For experiment 1 k^{asym} was set to 1, and for experiment 2 ten values were considered; five smaller than ones: k^{asym} =(0,1/5,1/4,1/3,1/2), and five greater: k^{asym} =(1,2,3,4,5,6).

7. Cell-cell adhesion: The strength of this rule is specified by p^{ADH}, a bio-mechanical property of the nodes that specify how strong per unit distance are the adhesion force between a pair of nodes (see Section 3.1). After a manual exploration of the parameter space, we found that the morphologies of simulated blastulae with p^{ADH} values between 10.0 and 20.0 were similar to that of real ones. Thus, we considered these values as the minimum and maximum cell-cell adhesion values. Intermediate values were also considered: p^{ADH}=(10.0, 15.0, 20.0) for the first experiment and p^{ADH}=(10.0, 12.0, 14.0, 16.0, 18.0, 20.0, 22.0, 24.0, 26.0) for the second experiment.

Other parameters: The other parameters in the SpiralMaker model were not changed since they do not have a large effect on spiralian development. The values for each of these parameters (see section 4) are biologically reasonable and allow cells to interact with other cells and perform changes in shape and position while keeping their physical integrity.

E. Bibliography:

Cather JN (1963) A time schedule of the meiotic and early mitotic stages of Ilyanassa. Caryologia 16(3):663-670.

Day S, Lawrence PA (2000) Measuring dimensions: the regulation of size and shape. *Development* 127(14):2977-2987.

Delile J, Doursat R, Peyriéras N (2013) Computational modeling and simulation of animal early embryogenesis with the MecaGen platform. *Comp Syst Biol* 2:359-405.

Freeman G, Lundelius JW (1982). The developmental genetics of dextrality and sinistrality in the gastropod Lymnaea peregra. *Roux Arch Dev Biol* 191(2):69-83.

Gilbert SF, Raunio AM (1997) Embryology: Constructing the organism. Sunderland, MA: Sinauer Associates.

Gillies TE, Cabernard C (2011) Cell division orientation in animals. *Curr Biol* 21(15):R599-609.

Goulding MQ (2009) Cell lineage of the Ilyanassa embryo: evolutionary acceleration of regional differentiation during early development. *PloS ONE* 4(5):e5506.

Grande, C. and Patel, N.H. (2008). Nodal signalling is involved in left–right asymmetry in snails. *Nature* 457, 1007-1011.

Graner F, Glazier JA (1992) Simulation of biological cell sorting using a two-dimensional extended Potts model. *Phys Rev Lett* 69(13):2013-2016.

Green JL, Inoue T, Sternberg PW (2008) Opposing Wnt pathways orient cell polarity during organogenesis. *Cell* 134(4):646-656.

Henry JQ. Perry KJ, Martindale MQ. 2006. Cell specification and the role of the polar lobe in the gastropod mollusc, *Crepidula fornicata*. *Dev Biol* 297:295–307.

Kajita, A., Yamamura, M. and Kohara, Y. (2003). Computer simulation of the cellular arrangement using physical model in early cleavage of the nematode Caenorhabditis elegans. *Bioinformatics*. 19, 704-716.

Kuroda R, Endo B, Abe M, Shimizu M (2009) Chiral blastomere arrangement dictates zygotic left–right asymmetry pathway in snails. *Nature* 462(7274):790-794.

Lu MS, Johnston CA (2013) Molecular pathways regulating mitotic spindle orientation in animal cells. *Development* 140(9):1843-1856.

Marin-Riera M, Brun-Usan M, Zimm R, Valikangas T, Salazar-Ciudad I (2015) Computational modelling of development by epithelia, mesenchyme and their interactions: A unified model. *Bioinformatics*. Btv527.

Maslakova SA, Martindale MQ, Norenburg JL. (2004) Fundamental properties of the spiralian developmental program are displayed by the basal nemertean Carinoma tremaphoros

(Palaeonemerta, Nemertea). *Dev Biol* 267(2):342-360.

Merkel J, Wollesen T, Lieb B, Wanninger A (2012) Spiral cleavage and early embryology of a loxosomatid entoproct and the usefulness of spiralian apical cross patterns for phylogenetic inferences. BMC Dev Biol 12(1):11.

Minc N, Burgess D, Chang F (2011) Influence of cell geometry on division-plane positioning. *Cell* 144(3):414-426.

Minc N, Piel M (2012) Predicting division plane position and orientation. *Trends Cell Biol* 22(4):193-200.

Newell GE (1948) A contribution to our knowledge of the life history of Arenicola marina L. *J Mar Biol Ass UK* 27(03):554-580.

Newman, T.J. (2005). Modeling multi-cellular systems using sub-cellular elements. *Math. Biosci. Eng.* 2, 611-622.

Purcell RM (1997) Life at low Reynolds number. *Am J Phys* 45(1):3-11.

Raven CP (1967) The distribution of special cytoplasmic differentiations of the egg during early cleavage in Limnaea stagnalis. *Dev Biol* 16(5):407-437.

Ren X, Weisblat DA (2006) Asymmetrization of first cleavage by transient disassembly of one spindle pole aster in the leech Helobdella robusta. *Dev Biol* 292(1):03-115.

Robert H (1902) Embryologie des Troques. *Arch de Zool Exper et Gen* 3:18-538.

Salazar-Ciudad I, Jernvall J, Newman SA (2003) Mechanisms of pattern formation in development and evolution. *Development* 130(10):2027-2037.

Sandersius, S.A. and Newman, T.J. (2008) Modeling cell rheology with the subcellular element model. *Phys. Biol.* 5, 015002.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature methods* 9(7): 671-675.

Song C, Wang P, Makse HA. 2008. A phase diagram for jammed matter. *Nature* 453(7195): 629-632.

Tee YH et al. (2015) Cellular chirality arising from the self-organization of the actin cytoskeleton. Nat Cell Biol 17(4):445-457.

Théry M, Bornens M (2006) Cell shape and cell division. *Curr Opin Cell Biol* 18(6):648-657.

Wilson EB (1892) The cell lineage of Nereis. A contribution to the cytogeny of the annelid body. *J Morph* 6:361-480.

Thesis discussion:

In this thesis work, we have assessed which are the cell processes involved in the different metazoan cleavage patterns and how are they involved in their morphological variation. After a thorough bibliographic research, we conclude that most metazoan cleavage patterns seem to use the same set of basic cell processes (which are in general evolutionarily ancient).

We have explored in more detail the spiral cleavage pattern by means of a specific computational model (the SpiralMaker) built from a general model of development (the EmbryoMaker). By using the SpiralMaker, we have shown that the spiral cleavage arises mainly due to the combination of few, and previously known, cell processes, namely: animal-vegetal polarization of cell division, Sachs' rule, cortical rotation and cell adhesion. These results have allowed us to discard other processes that were previously suggested to be involved in the spiral cleavage (Meshcheryakov, 1978; Takahasi and Shimizu, 1997; Aw and Levin, 2009).

Furthermore, the results drawn on the chapter 2 show that variations in the parameter values of these few cell processes can account for most of the morphological variation observed in the spiralian cleavage pattern. By using two alternative comparative methods, seven species-specific patterns of real organisms have been found within the theoretical morphospace generated by the model (these are the species for which we had experimental data). The patterns reminiscent of those of those species (four representatives of the main gastropods groups: *Crepidula*, *Planorbella*, *Lottia* and *Trochus;* a nemertean (ribbon worm), *Carinoma* and two polychaete worms *Nereis* and *Arenicola*.), arise for particular parameter combinations. Thus, the range of morphological variation that has been generated with our model is comparable to the range of variation in the cleavage patterns found between different spiralian taxa. This seems to be informative about the developmental processes underlying morphological differences at the level of blastulae between spiralian species.

Not all the different cleavage patterns generated by the model are equally likely to arise nor they are evenly distributed in the parameter space. Some of them (e.g. very regular patterns such as the one reminiscent of *Trochus*), appear under many different combinations of parameter values (they are found in a large volume of the parameter space) while others (e.g. patterns reminiscent of *Arenicola*) require a much more restricted combination of parameter values. In addition, the phylogenetic distance between the different species and the distance in the parameter space of the cleavage patterns of those species have no evident correlation. The possible generalization of our results to other non-spiralian patterns is discussed below.

These results could not be obtained without a computational model which takes into account

the main epigenetic factors involved in early development, and the basic fact that cells are physical entities subject to mechanical laws (e.g. deformation under different forces, adhesion between them, physical exclusion in space …). Few of the previously available models were readily suitable for implementing are the cell processes that are displayed in the spiral cleavage. In this sense, the SpiralMaker, an application of the more general EmbryoMaker to the specific developmental system of early cleavage, has proven successful for this purpose. In addition, the SpiralMaker is realistic enough to allow comparisons between the simulated and real blastulae (by two alternative methods specially developed for this).

From a more general perspective, the findings presented in this thesis could only be reliable under a developmental (Evo-devo) perspective, because as the model suggests, the different cleavage patterns (and in general the different morphologies found in organisms) can not be directly ascribed to particular genes (which have a limited causative role). Instead, the different patterns arise from an interplay between genes (which determine the developmental parameters) and the epigenetic factors (see introduction and Annex 1) in a highly non-linear way. This interplay imposes an internal logic structure on development, which results in a limited number of final patterns, each one with its own likelihood.

The spiral cleavage arises for a particular combination of cell processes.

The main result of this work is that the only combination of cell processes for which the spiral pattern arises are animal-vegetal polarization of cell division, Sachs' rule, cortical rotation and cell adhesion. Specifically, the animal-vegetal polarization is strictly required for cells to divide by primarily along the animal-vegetal axis (Raven, 1967; Gilbert and Raunio, 1997). This animal-vegetal polarization, in turn, depends on a molecular animal-vegetal gradient, what is known to be a common feature among metazoan oocytes (see chapter 1). The orientation of cell divisions along the animal-vegetal axis keeps the new arising groups of four cells (quartets) in the same relative position along this axis, and also keeps the overall tetra-radial symmetry of the blastula, with the four quadrants (the lineage of each blastomere in the 4-cell stage) in a close spatial relationship.

Cortical rotation is important in the third division to produce a coherent twist or rotation of all micromeres respect to the macromeres (Meshcheryakov and Beloussov, 1975; Henley, 2012). Our results have allowed us to discard other alternative hypotheses devoted to explain this symmetry breaking event (Freeman and Lundelius, 1975, Aw and Levin, 2009). However, as evidence suggest, some secondary mechanisms can be superimposed to cortical to cortical rotation, at least in some groups (Shibazaki et al., 2004, Wandelt and Nagy, 2004).

Sachs' rule is responsible of the left-right alternation after the 8-cell stage, which is in agreement with previous proposals (Guerrier, 1973). Finally, cell adhesion is required in order to keep the physical integrity of the blastulae (otherwise the blastulae could disaggregate by the spreading out of blastomeres). Our results suggest that the formation of the spiral cleavage pattern may not require signaling between cells, at least up to the 32-cell stage (this cell-cell signaling, however, is known to be crucial in cell fate determination (Grande and Patel, 2008)).

Our exhaustive combination of the different cell processes have also allow us to see that many combinations lead to other non-spiralian patterns. Amongst them, most (about 60%) lead to blastulae that lack the radial symmetry around the A-V axis that is characteristic of many metazoan blastulae (see chapter 1). Those asymmetric cleavage patterns are found when the Hertwig's rule (the cell division occurs along the longest axis of the cell) is applied, even when it is combined with other rules specifying the direction of cell division. Thus, we propose that Hertwig's rule can not be involved in the emergence of the spiral pattern, what is in contrast with some previous proposals (Meshcheryakov, 1978). This is likely due to the fact that is too sensitive to subtle changes in cell shape (that is, too noisy) as to determine the direction of cell division in a precise manner. However, these irregular blastulae in which each blastomere divides in an independent direction of the others (also known as chaotic cleavage), look similar to those blastula found in some cnidarians and sponges (see chapter 1).

The same holds when the cell divisions are oriented by cell-cell contacts. When this rule is implemented, alone or in combination with other processes, the spiral pattern is never found in simulations. Instead, this cell-cell contact rule results in radialized blastulae whose blastomeres are not in close contact between them, thus preventing the mechanical interactions between them that are crucial for other cell processes (e.g. cortical rotation) to act. These open configurations, however, resembles the cleavage pattern of some non-spiralian groups (e.g. ctenophores, see chapter 1).

Whether these rules (Hertwig's and cell-cell contact rule) that do not lead to spiral cleavage are actually involved or not in the cleavage pattern of these non-spiralian groups remains to be explored in further research.

Morphological variation within spiralian cleavage patterns can be explained by changes in the strength of the underlying developmental parameters:

The second main result of this thesis is that the morphological variation observed in the spiral cleavage pattern can be explained by quantitative variations in the relative strength of each of the few cell processes involved (plus asymmetric cell division, that is known to be an important driver of morphological variation within spiralian cleavage patterns (Merkel et al., 2012)). That is, species specific patterns arise for particular parameter combinations.

We arrived to this result by systematically varying the developmental parameters of the model (increments of 20% from the minimum to the maximum possible values in the strength of each cell process that we found to lead to spiral patterns). That way, a theoretical (generative) morphospace of all possible spiral patterns was built, and we searched within this morphospace simulated patterns that closely resemble those of seven real species. In order to do so, we developed a quantitative method that takes into account the relative contacts between adjacent blastomeres (this is which blastomeres, defined by their cell lineages, are in physical contact with each others).

This shows that the patterns reminiscent of the seven species considered appear in different points of our theoretical morphospace, suggesting that the morphological differences in the different spiral patterns arise from differences in the underlying developmental parameters (see Annex 2 for a detailed description of the parameter values that produce each species blastula).

The distribution of the different spiral pattern morphologies within the theoretical (generative) morphospace.

The comparative methods depicted in the previous paragraphs were also useful to assess how the different types of patterns (those alike but not necessarily identical to the pattern of each real species) are distributed within the parameter space. In order to do this, we defined a threshold value for similarity to each species, so that the regions of the parameter space that gave rise to blastulae whose similarity with the one of a given species was greater than the threshold formed a cloud (or hypervolume) within the parameter space. When the hypervolumes for different species were compared, results show that they differ in size. This suggests that, among all possible spiralian patterns (more than 6000 were simulated), some of them (those occupying larger hypervolumes) were more likely to appear than others. These easy-to-arise patterns would correspond to the generic forms described in the introduction, that is, stable patterns that do not require a tight genetic control to appear, but arise almost spontaneously due to the action of a few epigenetic factors

(Goodwin, 1996; Newman and Bhat, 2008).

Another feature of the hypervolumes of the different species is that they also differ in shape, and, in some species, they appeared as unconnected clouds. These results may indicate that, for each type of cleavage pattern (alike to the one of a real species), the cell processes that accounts for its morphological variation can be different. For instance, cortical rotation seems to account for most of the morphological variation in the *Nereis*-like blastulae (see Fig. 4B in Chapter 2).

Taken together, these results suggest that the morphological changes in the cleavage patterns are not tightly correlated with the progressive changes in the underlying developmental parameters. If we assume, as it seems reasonable, that these developmental parameters are controlled by genes (and thus their variation is due to random genetic mutations), this would imply a complex genotype-phenotype map (GPM).

SpiralMaker, a specific application of the general modeling framework Embryomaker, is perfectly suitable for modeling the cleavage process.

Overall, our results suggest that the spiral pattern can not be produced unless blastomeres have a specific positions in space in particular cell stages. In the transition between 4 to 8-cell stage, micromeres have to appear in the top (animal-most) part and they have to be in closer contact with each other in order for cortical rotation to be effective. In the 8 and ensuing cell stages, pairs of sister cells have to lie obliquely respect to the animal-vegetal axis, in order for Sachs' rule to produce the right-left alternation. In other words, the emergence of each developmental stage in the spiral pattern depends not only on the set of cell processes previously described but also on the spatial distribution of cells and molecules in the previous stage. This is extensive to the 1-cell (oocyte) stage, which has to be at least one animal vegetal gradient with polarizing activity in order for cell divisions to take place along this animal-vegetal axis (a second animal-vegetal gradient, responsible of asymmetric cell division, is also required in some spiralian groups for regulating the relative size between micro- and macromeres, despite this is nos strictly required for the spiral pattern to appear, as our results reveal). In addition, as explained in chapter 1, cortical rotation is known to arise as a consequence of the collective (self-organizing) behavior of flowing molecules of F-actin having chiral properties (Baum, 2006; Pinot et al., 2012; Tee et al., 2015).

All these facts point to epigenetic factors (mechanical interaction of cells in space, collective (self-organizing) behavior of chiral molecules, spatial regulatory gradients, distribution of cells in space …) as a crucial elements in the generation of the spiral cleavage. This is in agreement with

other empirical works in which some of these epigenetic factors are altered (subtle changes in blastomere position, inhibition of cortical rotation, egg-shell removal), being the further stages of the spiral pattern severely modified or even disrupted (Freeman and Lundelius, 1975; Kuroda et al., 2009; Shibazaki et al., 2004). In other words, the spiral pattern can not be totally understood if one is only concerned with genes (e.g. genes are known to be involved for instance in selecting left-right chirality in snails, though this genetically determined chirality can be switched simply by pushing micromeres (with a micro-needle) to their adjacent furrow between macromeres in the 8-cell stage).

Crucial as they are, epigenetic factors should be properly implemented to any computational model of development devoted to reproduce the early metazoan development (see chapter 1). In that sense, the implementation of these epigenetic factors (or their immediate effects) in the SpiralMaker is precise enough to generate cleavage patterns comparable to those of real species (see chapter 2). This may be partially explained by the fact that the SpiralMaker is a SEM-based model, which is able to reproduce both realistic cell shapes (and changes in cell shapes, through a cytoskeletal-like dynamics) and the realistic transmission of forces (arising from mechanical interactions) through cells and cell aggregates (Newman, 2004; Sandersius and Newman, 2008) (see Introduction).

Furthermore, our work shows that EmbryoMaker (the general modeling framework from which SpiralMaker is derived) is able to successfully model specific developmental systems such as the spiral cleavage considered in here. This represents another contribution to an growing list of specific developmental models built from this general platform: tooth development (Marin-Riera et al., in preparation), embryoids (Zimm et al., in preparation), and others. This shows the potential advantages of using general, readily available and user-friendly modeling framework instead of the classic organ-specific models.

Generalization of our results.

Our bibliographic research shows that most metazoan cleavage patterns use the same set of basic cell processes. These cell processes are also found in unicellular organisms and are thought to be evolutionarily old (Newman and Comper, 1990; Sebé-Pedros et al., 2010). The four rules that we have found to be responsible of the spiral cleavage pattern are a subset of these processes, and thus other, non-spiralian patterns are generated by other combinations of cell processes. For example, the blastomere arrangement of some phyla such as *Gnathostomulida* or *Acanthocephala* has been suggested to be explainable by the compressive effect of an outer coverage or eggshell (see Fig. 1 in Chapter 1). In other instances, such as in many arthropods, taxon-specific patterns seem to result mainly from the effect of abundant and dense yolk.

Our bibliographic research also shows that the distribution of cleavage patterns across metazoan taxa exhibits a weak phylogenetic relationship, and that some cleavage patterns are more widespread in the different metazoan taxa than others. These widespread patterns (e.g. spiral, spiral-like, pseudospiral …) have arisen many times during metazoan evolution, and it seems plausible they do this because they are more likely to arise from development (they require only basic, generic cell processes and weak genetic control over them).

These properties (the weak relationship between patterns and phylogeny and the differential likelihood between different patterns) are similar to those we have obtained for the specific case of the spiral cleavage pattern, and are known to be a general feature of most developmental systems (Alberch, 1982; Schuster et al., 1994; Salazar-Ciudad, 2001; Salazar-Ciudad and Jernvall, 2004). Taking this into account, and the fact that all metazoans (spiralians and non-spiralians) use common mechanisms during cleavage, we propose that our findings (that different spiral cleavage patterns appear due to variations in the strength of a set of simple cell processes, and that the different spiral patterns appear in the parameter space with a determined, but not strictly phylogenetic, distribution) are not exclusive of spiralians, but (at least in early stages) they may be extensive to other, non-spiralian groups.

In addition, the general trends relative to cleavage patterns drawn from chapter 1 allow us to offer a tentative explanation for the great abundance of the spiral pattern at high taxonomic levels. This abundance may be due to three, non exclusive reasons:

First, the results of our model show that the spiral pattern does not require too complex processes or many processes acting at the same time, but a rather low number of simple ones. In addition, these cell processes and their associated epigenetic factors, are known to be ancient for metazoans (Goodwin, 1996; Newman and Bhat, 2008). Thus, the spiral cleavage pattern is likely oneof the easiest-to-arise ones. Some non-spiral patterns, such as the radial one, are also rather common at high taxonomical levels, and thus may also be easy-to-arise ones (this possibility should be adressed by a specific computational model devoted to the radial pattern).

Second, the spiral pattern is a mechanically stable one, because by maximizing the cell-cell contact surface between adjacent cells, each micromere is forced to occupy the furrow between its two underlying macromeres. In this configuration the possible movements of the micromere are strictly limited by these adjacent macromeres, thus favouring that the relative neighborhood of each cell remains constant over time. The structural stability of the spiral pattern is also supported by the fact that many non-spiralian taxa with a weak genetic control of the cleavage process often exhibit transitory spiral-like (pseudospiral) patterns, that according to many authors (see chapter 1) is merely due to the optimal (less energy) packing configuration of small cell clusters having cell

adhesion.

Third, the aforementioned stability, along with the fact that spiral cleavage exhibits clear symmetry axes since early stages, allows for the early use of inductive mechanisms for cell fate and body axes determination. This early cell-fate determination is thought to be adaptive for those taxa spawning large amounts of small and relatively yolk-free eggs that result in a fast-developing, planktotrophic and free-swimming larvae, that is the case for most of groups forming the *Spiralia* superclade (Wray, 2000; Salazar-Ciudad, 2010). The swimming larvae of these taxa are so tiny that their functional organs are composed of one or very few cells, which requires a very precise inductive events (usually short-ranged signalling) between these cells taking place in a constant blastulae architecture (Salazar-Ciudad, 2010). In these cases even the mispalcement of a single cell would lead to catastrophic changes in the larval morphology (this is not the case in other animal groups, such as many vertebrates, in which inductive events associated with organogenesis take place much later in development and involve large number of cells) (ibid).

Future directions

As we have seen, although our results are mainly concerned with spiral (and to some extent with spiral-like) cleavages, the fact that other non-spiralian groups make use of similar cell processess in their cleavages, points towards a generality of our results. However, although one is tempted to speculate that this is the case, further implementations of the model would be required to explore if this is actually the case. These further implementations would be as follows:

First, for the non-spiralian patterns that have arisen in our model, the morphological varaition should be explored in the same systematic way that we did for spiralian patterns. In addition, other cell processes other than those we considered in the chapter 2 should be implemented. These may be, for instance: yolk as inhibitor of cell divisions, lack of synchrony between cell divisions (so that not all blastomeres divide at the same time), the compressive effect of various kinds of eggshells being more or less elongated of having more or less stiffness, etc … (see chapter 1). This would allow the simulation of a great number of cleavage patterns, and would allow us to explore the evolutionary transitions about these patterns.

Second, we should devise other, more precise, quantitative ways to compare between different kinds of blastulae at morphological level. This may imply the acquisition of new kind of data coming from 3D or 4D microscopy imaging (e.g. z-stack confocal microscopy). Then the 3D positions cells and cell shapes may be compared between real and simulated blastulae by means of quantitative morphometric methods (e.g. euclidean distance between cells' centroids, shared volume

between homologous blastomeres, …).

Third, it would be desirable to include in the analysis data relative to more species. As we explained in the introduction and in chapter 1, cleavage patterns are a very suitable systems for Evo-devo research, but we lack detailed knowledge about the early cleavage stages of many animal groups (even at phylum level). This is because most studies(developmental, genetic and others) use only model species that are chosen because of practical criteria (fast reproduction, few resources consumption, ease of breeding and handling) rather than their intrinsic phylogenetic position. Morphological data (4-D imaging) for a large set of organisms representative of all metazoan groups (not only model-systems) would be ideal to help identify the main factors involved in the different cleavage patterns and evolutionary transitions between them (thus contributing to an integrating understanding of cleavage), as well as to test the predictable power of the model.

Summarizing, a generalization of the model presented in here would shed light on how the interplay between developmental dynamics (and the epigenetic factors involved) and natural selection has sculpted the early stages of animal development, which is in the very core of the Evo-devo perspective on biological evolution.

Conclusions

- Despite their morphological diversity, the cleavage patterns found in metazoans can be classified in a relatively small set of cleavage types according to the spatial arrangement of their constituent blastomeres.

- Most of the metazoans cleavage patterns seem to use the same basic set of basic cell processes and epigenetic factors, being many of these processes and factors ancient and evolutionarily conserved.

- The spiral cleavage pattern, at least until the 32-cell stage, arises mainly due to the combination of few cell processes: an animal-vegetal polarization of cell division, Sachs' rule, cortical rotation and adhesion.

- The action of those cell processes also requires the existence of epigenetic factors (mechanical interaction of cells in space, collective (self-organizing) behavior of chiral molecules, spatial regulatory gradients, distribution of cells in space …). Thus, epigenetic factors are necessary elements in the generation of the spiral cleavage (and most likely also of non-spiralians), which argues against considering genes as the only causative factors in explaining biological evolution.

- Species specific patterns, as those of the seven different spiralian species considered in here, arise for particular parameter combinations. Thus, the morphological variation observed in spiralian cleavage can be mainly explained by quantitative variations in these few cell processes (this is changes in the relative strength of each of them).

- Some of these species-specific patterns (those arising under a very large number of parameter combinations) are much more likely to appear than others. These easy-to-arise patterns would correspond to patterns that do not require a tight genetic control to appear, but arise almost by default as a consequence of the action of basic epigenetic factors.

- Morphological changes in the cleavage patterns do not seem to be tightly correlated with the progressive changes in the underlying developmental parameters, which would imply that even relatively simple systems (such as the basic spiral cleavage that involves only five cell processes) exhibit a complex genotype-phenotype map (GPM).

- These previous points are in fully agreement with the Evo-devo point of view about biological evolution, and since all metazoans use mechanisms similar to spiralians to build up their cleavage patterns, the previous results are likely generalizable to all metazoans.

- In that sense, SpiralMaker (an application of the more general EmbryoMaker devoted to the modelling of early cleavage) has been perfectly suitable for modeling the first stages of cleavage, at least the spiral one. This is because, unlike other computational models of development, it implements most epigenetic factors (or their immediate developmental effects) involved in early development, in a realistic and computationally tractable manner.

BIBLIOGRAPHY:

Alberch, P. 1982. Developmental constraints in evolutionary processes. In Evolution and development (ed. J.T. Bonner), p313–332 .Dahlem Konferenzen. Springer-Verlag: Heidelberg.

Akiyama, M., Tero, A. and Kobayashi, R. 2010. A mathematical model of cleavage. J. Theor. Biol. 264: 84-94.

Avery, O., MacLeod, C., and McCarty, M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Inductions of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III». J. Exp. Med. 79(2): 137-158.

Balfour, F.M. 1881. A Treatise on Comparative Embryology . Vol two. Macmillan, London.

Barton N.H., Turelli, M. 1989. Evolutionary quantitative genetics: how little do we know? Ann. Rev. Genet. 23:337–370.

Bateson, W. 1894. Materials for the Study of Variation: treated with especial regard to discontinuity in the origin of species. Macmillan, London. Repinted 1992. Johns Hopkins University Press, Baltimore.

Bezem, J.J. andRaven, C.P. 1975. Computer simulation of early embryonic development. *J. Theor. Biol. 54*(1): 47-61.

Brun-Usan, M., Marin-Riera, M., and Salazar-Ciudad, I. 2014. On the effect of phenotypic dimensionality on adaptation and optimality. J. Evol. Biol. 27(12): 2614-2628.

Chargaff, E. 1950. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. Experientia. 6: 201-209.

Corominas-Murtra, B., Goñi, J., Solé, R.V., and Rodríguez-Caso, C. 2013. On the origins of hierarchy in complex networks. Proc. Natl. Acad. Sci. USA. 110(33): 13316-13321.

Darwin, C. 1859. On the Origin of Species. John Murray, London.

Dawkins, R. 1976. The selfish gene. Oxford University Press.

De Beer, G.R. 1954. Revised edition. Embryos and Ancestors. Oxford, Oxford University Press.

Delile, J., Doursat, R., and Peyriéras, N. 2013. Computational modeling and simulation of animal early embryogenesis with the MecaGen platform. Comput. Syst. Biol. 2: 359-405.

Dobzhansky, T.G. 1937. Genetics and the Origin of Species (Vol. 11). Columbia University Press.

Fisher, R.A. 1930. The Genetical Theory of Natural Selection. Clarendon, Oxford.

Freeman, G. and Lundelius, J.W. 1982. The developmental genetics of dextrality and sinistrality in the gastropod Lymnaea peregra. Roux. Arch. Dev. Biol. 191: 69-83.

Gavrilets, S. and Dejong, G. 1993. Pleiotropic models of polygenic variation, stabilizing selection and epistasis. Genetics. 134(2): 609-625.

Gilbert, S.F. and Raunio, A.M. 1997. Embryology: Constructing the organism. Sunderland, MA. Sinauer Associates.

Gilbert, S.F. 2003. The morphogenesis of evolutionary developmental biology. Int. J. Dev. Biol. 47: 467-477.

Goldschmidt, R.B. 1940. The Material Basis of Evolution. Yale University Press, New Haven.

Goodwin, B.C., and Trainor, L.E.H. 1980. A field description of the cleavage process in embryogenesis. J. Theor. Biol. 85(4): 757-770.

Goodwin, B.C. 1994. How the leopard changed its spots: The evolution of complexity. Princeton University Press.

Gould, S.J. 1977. Ontogeny and Phylogeny. Harvard University Press, Cambridge.

Gould, S.J., and Eldredge, N. 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology 3(02): 115-151.

Grande, C. and Patel, N.H. 2008. Nodal signalling is involved in left–right asymmetry in snails. Nature 457: 1007-1011.

Gould S.J. and Lewontin R.C. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc. R. Soc. Lond. B. 205: 581-598.

Gould, S.J., and Lloyd, E.A. 1999. Individuality and adaptation across levels of selection: How shall we name and generalize the unit of Darwinism?. Proc. Natl. Acad. Sci. USA. 96(21): 11904-11909.

Graner, F. and Glazier, J.A. 1992. Simulation of biological cell sorting using a two-dimensional extended Potts model. Phys. Rev. Lett. 69(13): 2013–2016.

Haeckel, E. 1866. Generelle Morphologie der Organismen: Allegeneine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründendet durch die von Charles Darwin reformite Descendenz-Theorie. 2 Volumes. G. Reimer, Berlin. Vol. 1, pp. 43- 60

Haig D. 2004. The (dual) origin of epigenetics. Cold Spring Harb. Symp. Quant. Biol. 69:67-70.

Haldane, J.B.S. 1953. Foreword. In: R. Brown, and J. F. Danielli (eds.) Evolution. (SEB Symposium VII). Cambridge University Press. Cambridge. Pp. ix – xix.

Hall, B.K. 2000. Balfour, Garstang, and de Beer: The first century of evolutionary embryology. Amer. Zool. 40: 718-728.

Hansen, T.F. and Wagner, G.P. 2001. Modeling genetic architecture: a multilinear theory of gene interaction. Theor. Popul. Biol. 59(1): 61-86.

Hardy, H.G. 1908. Mendelian propotions in a mixed population. Science, 28: 49-50.

Hejnol, A. 2010. A twist in time- the evolution of spiral cleavage in the light of animal phylogeny. Integr. Comp. Biol. 50: 695-706.

Henley, C.L. 2012. Possible origins of macroscopic left-right asymmetry in organisms. J. Stat. Phys. 148: 741-775.

Henry, J.Q. 2014. Spiralian model systems. Int. J. Dev. Biol. 58: 389-401.

Hogeweg, P. 2000. Evolving mechanisms of morphogenesis: on the interplay between differential adhesion and cell differentiation. J. Theor. Biol. 203(4): 317-33.

Honda, H., Tanemura, M. and Nagai, T., 2004. A three-dimensional vertex dynamics cell model of space-filling polyhedra simulating cell behavior in a cell aggregate. J. Theor. Biol. 226(4): pp.439–453.

Honda, H., Motosugi, N., Nagai, T., Tanemura, M. and Hiiragi, T. 2008. Computer simulation of emerging asymmetry in the mouse blastocyst. Development 135: 1407-1014.

Horder, T.J. 1989. Syllabus for an embryological synthesis. In: Wake DB, Roth G, editors. Complex organismal functions. New York: John Wiley.

Jablonka, E., and Lamb, M.J. 2007. Précis of evolution in four dimensions. Behav. Brain. Sci. 30(4): 353-365.

Jacob, F. 1977. Evolution and tinkering.. Science 196: 1161-1166.

Jaeger, J., Blagov, M., Kosman, D., Kozlov, K.N., Manu, Myasnikova, E., Surkova, S., Vanario-Alonso, C.E., Samsonova, M., Sharp, D.H. and Reinitz, J. Dynamical analysis of regulatory interactions in the gap gene system of Drosophila melanogaster. Genetics. 2004(4): 1721-1737.

Jernvall, J. 2000. Linking development with generation of novelty in mammalian teeth. Proc. Natl. Acad. Sci. USA. 97: 2641-2645.

Kajita, A., Yamamura, M. and Kohara, Y. 2003. Computer simulation of the cellular arrangement using physical model in early cleavage of the nematode Caenorhabditis elegans. Bioinformatics. 19: 704-716.

Kaneko, K. 1990. Clustering, coding, switching, hierarchical ordering, and control in a network of chaotic elements. Physica D. 41(2): 137-172.

Kauffman, S. A. 1993. The Origins of Order. Oxford University Press, New York.

Klingenberg, C.P. and Leamy, L.J. 2001. Quantitative genetics of geometric shape in the mouse mandible. Evolution 55: 2342–2352.

Kowalevsky, A. 1866. Entwickelungsgeschichte der einfachen Ascidien. Mémoires de L'Académie Impériale des Sciences de St. -Pétersbourg , VII Série. Tome X: 1- 22.

Kropotkin, P. 2012. Mutual aid: A factor of evolution. Courier Corporation.

Kuroda, R., Endo, B., Abe, M. and Shimizu, M. 2009. Chiral blastomere arrangement dictates zygotic left–right asymmetry pathway in snails. Nature 462: 790-794.

Laumer, C.E., Bekkouche, N., Kerbl, A., Goetz, F., Neves, R.C., Sørensen, M.V., Kristensen, R.M., Hejnol, A., Dunn, C.W., Giribet, G. et al. 2015. Spiralian phylogeny informs the evolution of microscopic lineages. Curr. Biol. 25: 2000-2006.

Lillie, F.R. 1898. Adaptation in cleavage. In Biological Lectures from the Marine Biological Laboratory, Woods Hole, Massachusetts. Ginn, Boston, pp. 43–67.

Lourenço, J., Galtier, N., and Glémin, S. 2011. Complexity, pleiotropy, and the fitness effect of mutations. Evolution, 65(6): 1559-1571.

Løvtrup, S. 1978. On von Baerian and Haeckelian recapitulation. Syst. Biol. 27(3): 348-352.

Lynch, M., and Walsh, B. 1998. Genetics and analysis of quantitative traits (Vol. 1). Sunderland,

MA. Sinauer.

Marée, A.F., Grieneisen, V.A., and Hogeweg, P. 2007. The Cellular Potts Model and biophysical properties of cells, tissues and morphogenesis. In: Single-cell-based models in biology and medicine. Birkhäuser Basel. pp. 107-136.

Morgan, T.H. 1932. The Scientific Basis for Evolution . W.W. Norton, NY.

Salazar-Ciudad, I., and Marín-Riera, M. 2013. Adaptive dynamics under development-based genotype-phenotype maps. Nature 497(7449): 361-364.

Marin-Riera, M., Brun-Usan, M., Zimm, R., Valikangas, T. and Salazar-Ciudad, I. 2015. Computational modeling of development by epithelia, mesenchyme and their interactions: A unified model. Bioinformatics 32(2): 219-225.

Martin, G. and Lenormand, T. 2006. A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution 60: 893-907.

Maynard-Smith, J., Burian, R., Kauffman, S.A., Alberch, P., Campbell, J., Goodwin, B.C., Lande, R., Raup, D. and Wolpert, L. 1985. Developmental constrains and evolution. Quart. Rev. Biol. 60: 265-287.

McGhee, G.R. 2001. Exploring the spectrum of existent, nonexistent and impossible biological form. Trends. Ecol. Evol. 16(4): 172-173.

Meinhardt, H. 1982. Models of Biological Pattern Formation. London: Academic Press.

Meshcheryakov, V.N. and Beloussov, L.V. 1975. Asymmetrical rotations of blastomeres in early cleavage of gastropoda. Roux. Arch. Dev. Biol. 177: 193-203.

Minc, N. and Piel, M. 2012. Predicting division plane position and orientation. Trends. Cell. Biol. 22: 193-200.

Morgan, T.H. 1932. The Scientific Basis for Evolution. W.W. Norton, NY.

Newman, S.A. and Müller, G.B. 2000. Epigenetic mechanisms of character origination. J. Exp. Zool. 288(4): 304-317.

Newman S.A. and Comper, W.D. 1990. 'Generic' physical mechanisms of morphogenesis and pattern formation. Development. 110(1): 1-18.

Newman, T.J., 2005. Modeling multicellular systems using subcellular elements. Math. Biosci. Eng. 2(3): 613-624.

Newman, S.A., and Bhat, R. 2008. Dynamical patterning modules: physico-genetic determinants of morphological development and evolution. Phys. Biol. 5(1): 015008.

Newman, S.A. 2012. Physico-genetic determinants in the evolution of development. Science 338(6104): 217-219.

Nielsen, C. 1994. Larval and Adult Characters in Animal Phylogeny. Am. Zool. 34: 492-501.

Niklas, K.J. 1999. Evolutionary walks through a land plant morphospace. J. Exp. Bot. 50(330): 39-52.

Nirenberg, M., Leder, P., Bernfield, M., Brimacombe, R., Trupin, J., Rottman, F. and O'Neal, C. 1965. RNA codewords and protrin synthesis, VII. On the general nature of the RNA code. Proc. Natl. Acad. Sci. USA. 53(5): 1161-1168.

Nowak, M.A., Boerlijst, M.C., Cooke, J., et al. 1997. Evolution of genetic redundancy. Nature. 388(6638): 167-171.

Ospovat, D. 1981. The Development of Darwin’s Theory: Natural History, Natural Theology, and Natural Selection, 1838-1859 . Cambridge University Press, Cambridge.

Oster, G.F., Alberch, P. 1981. Evolution and bifurcation of developmental programs. Evolution 36: 444-459.

Osterfield, M. et al., 2013. Three-Dimensional Epithelial Morphogenesis in the Developing Drosophila Egg. Dev. Cell, 24(4): 400–410.

Paley, W. 1802. Natural Theology or, Evidence of the Existence and Attributes of the Deity. R. Faulder, London.

Pinho, R., Borenstein, E. and Feldman, M.W. 2012. Most Networks in Wagner's Model Are Cycling. Plos One 7(4): e34285.

Polly, P.D. 2005. Development and phenotypic correlations: the evolution of tooth shape in Sorex araneus. Evol. Dev. 7: 29–41.

Raup, D.M. 1961 The geometry of coiling in gastropods. Proc. Natl. Acad. Sci. U.SA. 47: 602-609.

Roux, W. 1894. The problems, methods and scope of developmental mechanics. In: Biological lectures of the Marine Biology Laboratory, Woods Hole . Ginn, Boston., pp. 149–190.

Salazar-Ciudad, I., Garcia-Fernandez, J., and Sole, R.V. 2000. Gene networks capable of pattern formation: from induction to reaction- diffusion. J. Theor. Biol. 205: 587-603.

Salazar-Ciudad, I., Solé, R.V., Newman, S.A. 2001. Phenotypic and dynamical transitions in model genetic networks. II. Application to the evolution of segmentation mechanisms. Evol. Dev. 3(2): 95-103.

Salazar-Ciudad, I., Jernvall, J. and Newman, S.A. 2003. Mechanisms of pattern formation in development and evolution. Development 130: 2027-2037.

Salazar-Ciudad, I. 2006. Developmental constraints vs. variational properties: How pattern formation can help to understand evolution and development. J. Exp. Zool. B. Mol. Dev. Evol. 306(2): 107-125.

Salazar-Ciudad, I. 2007. On the origins of morphological variation, canalization, robustness, and evolvability. Int. Comp. Biol. 47(3): 390-400.

Salazar-Ciudad, I., 2008. Tooth morphogenesis in vivo, in vitro, and in silico. Curr. Top. Dev. Biol.

81: 341-371.

Salazar-Ciudad, I. and Jernvall, J., 2010. A computational model of teeth and the developmental origins of morphological variation. Nature 464(7288): 583-586.

Sandersius, S.A. and Newman, T.J. 2008. Modeling cell rheology with the subcellular element model. Phys. Biol. 5: 015002.

Schuster, P., Fontana, W., Stadler, P.F., and Hofacker, I.L. 1994. From sequences to shapes and back: a case study in RNA secondary structures.Proc. Roy. Soc. B: Biol. Sci. 255(1344): 279-284.

Stomps, T.J. 1954. On the rediscovery of Mendel's work by Hugo de Vries. J. Hered. 45(6): 293-294.

Thom, R.1977. Stabilité structurelle et morphogénèse, Interédition, París.

Todes, D. 1989. Darwin without Malthus: The Struggle for Existence in Russian Evolutionary Thought. Oxford, New York.

Valentine, J.W. 1997. Cleavage patterns and the topology of the metazoan tree of life. Proc.Natl. Acad. Sci. USA. 94(15): 8001-8005.

Waddington, C.H. 1942. Canalization of development and the inheritance of acquired characters. Nature 150: 563-565.

Waddington, C.H. 1953. Epigenetics and evolution. In R. Brown and J. F. Danielli (eds). Evolution. (SEB Symposium VII). Pp. 186-199. Cambridge University Press. Cambridge.

Wagner, A. 1994. Evolution of gene networks by gene duplications: a mathematical model and its implications on genome organization. Proc. Natl. Acad. Sci. USA. 91(10): 4387-4391.

Wandelt, J. and Nagy, L.M. 2004. Left-right asymmetry: more than one way to coil a shell. Curr. Biol. 14: R654-656.

Watson, J. D., and Crick, F. H. 1953. Molecular structure of nucleic acids. Nature 171(4356): 737-

738.

Weissman, A. 1875. Über den Saison-Dimorphismus der Schmetterlinge. In: Studien zur Descendenz-Theorie. Engelmann, Leipzig

Wilson, E.B. 1898. Cell lineage and ancestral reminiscence. In Biological Lectures from the Marine Biological Laboratories, Woods Hole, Massachusetts . Ginn, Boston, pp. 21-42.

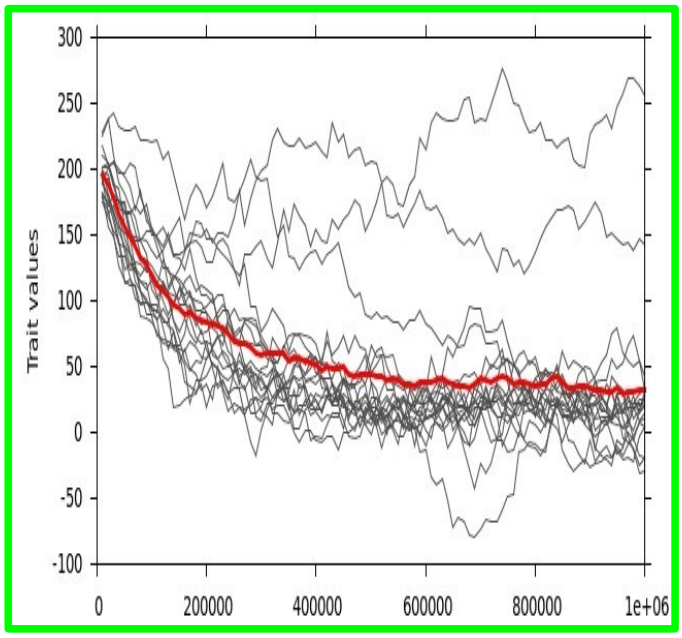
Wright, S. 1949. The genetical structure of populations. Ann. Eugenic. 15(1): 323-354.

Zammataro, L., Serini, G., Rowland, T., and Bussolino, F. 2007. Embryonic cleavage modeling as a computational approach to sphere packing problem. J. Theor. Biol. 245(1): 77-82.

ANNEX I

On the effect of phenotypic dimensionality on adaptation and optimality

The content of this annex is a paper published in 2014 by *Miguel* Brun-Usan, Miquel Marin-Riera and Isaac Salazar-Ciudad in *Journal of evolutionary biology (27*(12), 2614-2628).



On the effect of phenotypic dimensionality on adaptation and optimality

Authors: Miguel Brun-Usan [1], Miquel Marin-Riera [1], Isaac Salazar-Ciudad [1,2,*]

[1] Genomics, Bioinformatics and Evolution. Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Barcelona, Spain.

[2] Evo-devo Helsinki community, Centre of excellence in computational and experimental developmental biology, Institute of Biotechnology, University of Helsinki, PO Box 56, FIN-00014 Helsinki, Finland.

* Author to whom correspondence should be sent: isalazar@mappi.helsinki.fi
(Tel: +358-9-191 59352)

Running title : "Phenotypic dimensionality affects optimality"

Abstract:

What proportion of the traits of individuals has been optimally shaped by natural selection and what has not? Here we estimate the maximal number of those traits by using a mathematical model for natural selection in multi-trait organisms. The model represents the most ideal conditions for natural selection: a simple genotype-phenotype map and independent variation between traits. The model is also used to disentangle the influence of fitness functions and the number of traits, *n*, *per se* on the efficiency of natural selection. We also allow *n* to evolve.

Our simulations show that, for all fitness functions and even in the best conditions optimal phenotypes are rarely encountered, only for *n=1*, and that a large proportion of traits are always far from their optimum, specially for large *n*. This happens to different degrees depending on the fitness functions (additive linear, additive non-linear, Gaussian and multiplicative). The traits that arise earlier in evolution account for a larger proportion of the absolute fitness of individuals. Thus, complex phenotypes have, in proportion, more traits that are far from optimal and the closeness to the optimum correlates with the age of the trait.

Based on estimated population sizes, mutation rates and selection coefficients we provide an upper estimation of the number of traits that can become and remain adapted by direct natural selection.

Introduction:

The degree of fit of organisms to their environment is a long discussed subject in evolutionary biology (Parker & Maynard-Smith, 1990). Which proportion of the quantitatively describable phenotypic features (or traits) of an organism have values determined by natural selection for these precise values? Which proportion are simply a result of other processes? Before Darwin an influential view was that the phenotype is optimal in all its traits (Paley, 1802) .
In evolutionary biology it is natural selection that optimizes the phenotype (Maynard-Smith, 1978). Most single traits exhibit genetic variation and, thus, natural selection can act on them. Since selection has had the opportunity to act on most traits, it has been argued (Fisher, 1930) that most traits should be optimally adapted to their environment (Cain, 1964; Orzack & Sober, 1994).

Although natural selection is likely one of the most important factors in evolution, there are other widely recognized important factors that preclude organisms from being optimal (Haldane, 1927; Wright, 1932; Gould & Lewontin, 1979; Alberch, 1982; Kimura, 1983). Some authors stress the variational non-independence between traits, and that that precludes the simultaneous optimization of all traits (Bock, 1959; Lande & Arnold 1983). Then, many traits are not shaped by direct natural selection on them but by natural selection on traits with which they co-vary. These interdependences arise because of the developmental (Alberch, 1982; Salazar-Ciudad, 2006) and genetic processes by which phenotypic variation is produced and because of functional or ecological trade-offs between traits (Wake & Larson, 1987; Shoval *et al.*, 2012). Others (Salazar-Ciudad, 2007) stress that, while there is variation in most single traits, many combinations of trait values are not possible, or are rare, from the developmental processes by which phenotypic variation arises. Thus the direction of phenotypic evolution would be determined by a dialogue between which variation is developmentally possible and which of it is fit.

In this article we focus on how the dimensionality of the phenotype affects its optimality. By dimensionality we mean the number of independent quantitative phenotypic traits. Traits are understood in this article as any measurable phenotypic feature. However, our model considers only independent traits. Independent traits are those that can change, due to mutation, without affecting others traits. This is certainly an idealization since traits are independent from each other to different degrees (Alberch, 1982; Lande & Arnold, 1983). By optimality we refer to the proportion of the traits in an individual organism that are optimal. Different researchers understand the concept of trait and phenotype dimensionality in slightly different ways but there is a relatively rich literature devoted to understand the relationship between the dimensionality of a phenotype and its optimality.

Perhaps the most influential work on that topic is Fisher's geometric model (Fisher, 1930). In Fisher's geometric model the phenotype of an organism can be conceptualized as a point in a *n*-dimensional phenotypic space, where the dimensions of that space correspond to the traits of the organism. The fitness of an organism is, then, defined as the euclidean distance to the point in this space where the optimum is (this is to an optimal combination of values in each trait). Mutations are simply *n*-dimensional steps in this *n*-dimensional phenotypic space. In other words a mutation is a *n*-dimensional vector of change. There is thus universal pleiotropy (Orr, 2000) in the sense that every single mutation affects all the traits in an organism. Mutations that decrease the distance to the optimum are favored by natural selection. In a one dimensional phenotypic space (one trait space) half of the mutations are going to be, on average, adaptive since the optimal trait value is necessarily larger or smaller than the current trait value. In a two-dimensional phenotypic space (two traits space) one can represent the optimal phenotype as a point (for example at the origin of coordinates) and any other phenotype A as a point away from the origin in this space. All the mutants that decrease the distance of A to the optimum are going to be out of the circle centered at the origin and passing by point A . This implies than less than half of the mutants from A (the ones within the circle) are going to be adaptive. Extending this argument to *n*-dimensional spaces and hyper-spheres it can be seen that the frequency of adaptive mutations decreases with the dimensionality of the phenotype (Orr, 2000; Tenaillon *et al*,. 2007). This decreases the capacity of natural selection to lead the population to the optimum since the larger *n* the larger the proportion of mutations that are non-adaptive.

This argument has been implemented into a more accurate mathematical model by Orr (Orr, 2000) to calculate that the rate of adaptation decreases with the number of phenotypic traits in a phenotype as n^{-1}, where *n* is the number of traits. Other authors (Poon & Otto, 2000; Tenaillon *et al.*, 2007) have used modifications of this model to explore different evolutionary questions such as the distribution of mutation fitness effects (Martin & Lenormand, 2006; Lourenço *et al.*, 2011), evolution in a gradually changing environment (Matuszewski *et al.* 2014) and the evolution of phenotypic complexity as in here (Martin & Lenormand, 2006; Tenaillon *et al.*, 2007; Lourenço *et al.*, 2011). Some of those models provide an additional reason for expecting a decrease in adaptation for large *n* (Orr, 2000; Lourenço *et al*, 2011). They consider a Gaussian fitness function in which fitness is a negative exponential function of the euclidean [6] in here). Thus, the larger *n* the larger the number of terms contributing to the euclidean distance and thus the larger, on average, that distance can be. Since this distance is the exponent in the Gaussian distribution the larger this exponent is the more steep does the distribution decrease from the origin and the smaller is the variance in fitness effects. This is, the larger *n* the more mutations tend to be close to neutral and the

larger is the drift load and thus *n* decreases adaptation (Lourenço *et al*, 2011)

In here we extent the study of how *n* influences optimality by considering other fitness functions beyond the Gaussian function. As we will discuss these other fitness functions represent different ecological and functional plausible interdependences between the traits in each organism and have consequences for our understanding of which proportion of the traits of organisms are likely to be optimal. In general individuals with a very well adapted trait (for example a bird with the right beak size to eat the most abundant food) may not be well adapted in other traits (for example in its color, thermal isolation, etc...) and thus globally may have a low individual fitness. Since selection acts on individuals, and not on individual traits, and individuals are made of many different phenotypic traits the chances of an individual being selected depend on many of the traits in an organism. In other words, the fitness of an individual is a function (that we call in here the *fitness function* with the Gaussian fitness function being one example) of the values of its traits and both the form of this function and the number of those traits can be expected to have an effect on the dynamics of adaptation.

To explore the consequences of that argument and of the fitness functions on phenotypic evolution we built a very simple model of phenotypic evolution. For a neater understanding of the consequences of phenotypic dimensionality (without other confounding effects) it is more convenient to use a very simple model with a simple genotype-phenotype map, independent variation between traits and a non-changing environment. Such an over-simplistic ideal model is also convenient as an extreme null model of what to expect if natural selection would be the only relevant force or factor in determining evolutionary change (a simple genotype-phenotype map, no functional trade-offs, etc...).

There are, however, many other models that are much more realistic and consider pleiotropy and epistasis (Barton & Turelli, 1991; Kondrashov & Turelli, 1992; Zhivotovsky & Feldman, 1992; Wagner, 1994; Orr, 2000; Hansen & Wagner, 2001; Jones *et al.*, 2004; Hill *et al.*, 2006; Weinreich & Knies, 2013) and even development as such (Solé *et al.*, 2002; Salazar-Ciudad & Jernvall, 2010; Salazar-Ciudad & Marin-Riera, 2013). Some of this work provides important reasons (related to gene networks in development or to pleiotropy between gene effects) for the non-optimality of multi-trait phenotypes. Here we simply explore a different reason for that that can be understood without the complexities of the realistic genotype-phenotype maps or any of their aspects at different levels (pleiotropy, dominance, epistasis, etc...).

The literature in multi-locus evolution is also relevant in here (Kirkpatrick *et al.*, 2002; Gerrish *et al*., 2007). In multi-locus evolution a single phenotypic trait (or simply the individual fitness) is determined by a number of loci with different alleles. If there is strong linkage then it can

be that an adaptive allele in a locus cannot reach fixation because it is linked with other non-adaptive alleles in the rest of the loci. This is somehow analogous to the situation described in the case of selection in individuals with multi-trait phenotypes. In this respect the questions presented before would be translated to the number of linked loci that can be maintained as optimal by natural selection and the fitness functions would relate to how different loci interact with each other to lead to a specific individual fitness. However, as this article studies only phenotypic evolution (and not the evolution of its underlying genetics), our model uses some assumptions not present in multi-locus models while discarding others present in those models. Mutation rates per trait should be expected to be larger than mutation rates per locus since each trait is usually affected by many loci. Our model then does not consider mutation rates on loci but directly on phenotypic traits. Thus, our model makes no assumptions and is not concerned about how genes interact to produce the phenotype (there are no development, gene networks, pleiotropy, episistasis, dominance, loci or alleles as such). Our model simply assumes, very simplistically and largely unrealistically, that the probability of change is the same for all traits and that traits change independently from each other. Then we will not ask about how a specific set of alleles replace each other (as in many multi-locus models) but about the long-term evolution of these continuous phenotypic traits themselves (while multi-locus models are generally devoted to study the kinetics of replacement between existing alleles and are thus relatively short-term evolution and not focused on the phenotype).

There are also models devoted to study some aspects of selection on multi-trait phenotypes (most notably Lande & Arnold, 1983). Lande's and Arnold's multivariate analog of breeders equation (Lande & Arnold, 1983) is used to calculate how selection in one (or multiple) trait(s) lead to selection responses on other traits because of trait co-variations. Population size (and thus drift), mutation rates and the number of traits *per se* do not explicitly enter on Lande's equations. Since we want to study the effect of those we do not use Lande's model (although our model is not in contradiction with it).

Our article does not focus on the rate of adaptation, as for example in (Orr, 2000), but on optimality, this is on which proportion of the traits of a phenotype attain phenotypic values close to their optimal values in evolution, and on the distribution of this optimality through the traits in a phenotype and over evolutionary time. This is, which proportion of the traits of a phenotype are close to their optimum (and which proportion is not) and how that distribution changes with n and with the fitness functions. In addition, some of the simulations we perform on the model consider that n itself can change by mutation and, thus, considers, in a simplistic way, how phenotypic complexity can evolve. Then we study how the distribution of optimality changes over evolutionary time and how, and to which extent, the traits that arise earlier in evolution are more likely to be

optimal.

The mutable *n* model provides also a simple null-model on how phenotypic complexity, understood here as the number of traits, could evolve for a given mutation rate and populational size. This is a null model because it does not consider the more realistic limitations induced by trait co-variation, the mechanisms by which the phenotype is produced (e.g. development) and many other factors (dominance, pleiotropy, etc...). This means that the lack of adaptation found in our model can not be due to any genetic, developmental or functional trade-offs, but arises once all these well known limitations have been removed. Our model can also be seen as an extreme or upper boundary case for the rest of models. In it everything is set to allow natural selection, for a given set of selection coefficients, to be most efficient in changing the phenotype. Thus, for example, the estimation of the number of independent traits that can be maintained optimal by natural selection is the upper limit, any other model would provide lower estimations (because of a more realistic genotype-phenotype map or pleiotropy or something else). In the discussion we analyze how these other factors relate to our results.

Materials and methods:

Individuals have a given number of traits, *n*, and populations a given population size, *P*, and mutation rate, μ. The population is evolving towards a specific optimum characterized by specific values in each trait. Traits can vary by mutation and each trait variation is independent from that of the others.

Each individual phenotype is defined by a vector of *n* continuous quantitative phenotypic traits. An individual fitness is a function of the distance of each trait to that trait's optimal value:

$$d_{ij}=|t_{ij}-t_{oj}| \qquad [1]$$

Where t_{ij} is the value of trait *j* in individual *i*,and t_{oj} is the optimal value of trait *j*. Thus, the optimal phenotye is the one in which each trait is equal to its optimal value. For simplicity and to ensure the most ideal conditions for the action of natural selection we consider that this optimum does not change over time (even if in general optimal phenotypes should change over time with ecology, Bell, 2010). In an individual, however, some of the traits may reach, in the simulation of evolution, their optimal value while others may never do. In this case the individual is not optimal but some of their traits are. To measure to which extent this is the case we define a quantity, the optimality of an individual *i*, as $O_i=W_i/n$, where W_i is the absolute fitness of individual *i*. When

O_i<1 the individual *i* is sub-optimal, when $t_{ij}=t_{oj}$ the trait *j* in individual *i* is considered optimal.

Taking into account that there exist a long-lasting debate about what a trait actually is and how many traits do organisms have, (Violle *et al.*, 2007) we adopt a minimal definition of trait: any measurable phenotypic feature of an individual. Thus, in the model all traits of an individual are functional by default, and no neutral traits are introduced *ad hoc*. However, for the reasons mentioned before, we only consider traits that vary independently. Five different fitness functions are used:

-Additive inverse: $$W_i=\sum_{j=1}^{n}\frac{s_j}{1+d_{ij}} \quad [2]$$

-Additive linear: $$W_i=\Phi\left[\sum_{j=1}^{n}s_j\left(1-\frac{d_{ij}}{d_{max}}\right)\right] \quad [3]$$

-Additive linear with cutoff: $$W_i=\sum_{j=1}^{n}s_j\Phi\left[1-\frac{d_{ij}}{d_{max}}\right] \quad [4]$$

-Additive exponential: $$W_i=\sum_{j=1}^{n}s_j e^{-d_{ij}} \quad [5]$$

-Gaussian fitness function: $$W_i=\ e^{\frac{-z^2}{2}}\quad where\quad z=\sqrt{\sum_{j=1}^{n}\left(s_j d_{ij}\right)^2} \quad [6]$$

-Multiplicative: $$W_i=\prod_{j=1}^{n}\left(\frac{1}{1+d_{ij}}\right)^{\frac{s_j}{n}} \quad [7]$$

Where W_i is the absolute fitness of individual *i*, d_{max} is the inverse slope of the linear fitness function, s_j is the selection coefficient of trait *j* and Φ is a cutoff function (it is 0 if its argument is smaller than 0 and equal to the argument if the argument is larger or equal to 0). Each term in each fitness function is what we call the fitness contribution of trait *j* on individual's *i* absolute fitness or simply f_{ij}.

The additive fitness functions (Eqs. [2] to [5]) represent the situation in which each trait contributes to its bearer's fitness independently from the rest of its traits. In the inverse fitness function [2] the fitness contribution of a trait is inversely proportional to the distance to that trait optimum. One is added to the denominator to restrict that contribution to a maximum value of 1. In the inverse [2] and exponential [5] fitness functions the fitness contribution of a trait decreases non-linearly with d_{ij}. These two equations represent the situation in which when a trait is close to the optimum (for example a beak approaching the size of one of the flowers tube from which a bird

feeds) further improvements in fitness are easily accomplished by changes in a trait. However, when the trait is far from the optimum (for example when the beak is so large that it does not fit at all in the flower tube) then further changes (increases or decreases) in the trait are unlikely to lead to significant fitness chances. This situation is not considered, for example, in the fitness functions considered in Orr's work, although it is likely common in the ecology of organisms. The linear cutoff fitness function is similar in the sense that beyond d_{max} further increases in d_{ij} have no effect on fitness. After this point this fitness function is effectively flat (constant). In that sense all the additive functions but the linear are very similar when a trait is very far away from the optimum. Note that while in equation [4] the cutoff applies to every trait (and thus the contribution of a trait is never negative) in equation [3] the cutoff applies only to the sum of the contributions of all the traits and thus traits may contribute negatively to individual fitness (but the cutoff precludes individual fitness to ever be negative).

In the additive linear fitness function [3] it is not the case and traits that are very far away ($d_{ij}>d_{max}$) from their optimum contribute negatively to the fitness of its bearer. This is a less realistic scenario in which even if a trait is very far away from its optimum any change in the direction of the optimum would always imply a fitness increase (and always in the same proportion).

This occurs also in the Gaussian fitness function [6]. In it z is the euclidean distance to the optimum. This is the fitness function considered in Orr's and related work(Orr 2000; Martin & Lenormand, 2006; Lourenço *et al.*, 2011). This fitness function [6] is a bit more complicated because when most traits are very close or very far away from the optimum the slope of the functions is small and it is large at intermediate distances. In addition, if any of the traits is very far away from its optimum the overall fitness of the individual is very small (irrespective of how optimal are the other traits). This happens, in contrast with the linear fitness function, in a non-linear way so that if any trait is far away from its optimum the whole individual is really maladaptive. In this fitness function, thus, there is an implicit effective dependence between traits in the sense that the contribution of a trait to the individual fitness would be very small if any of the traits in the individual is very far away from the optimum.

In the multiplicative fitness function [7] individual fitness arises from the cooperative contribution of all traits. In this case two traits that are very far away from the optimum can largely decrease an individual fitness (in the extreme any trait with f_{ij} equal to 0 would lead to an individual with 0 fitness). The product implies that, for the same average distance to the optimum per trait, maximal fitness is reached when all traits are around the same distance to their optimum. In this sense an individual reaches high fitness if it has roughly equally high adaptation in all its traits. This represent the situation, as in the Gaussian fitness function, in which having a very adapted trait (for

example a beak of the most optimal size) is useless if some other trait is poorly adapted (for example if the bird legs are so small that the bird can not walk nor land).

These fitness functions represent a simple tractable subset of all the possible ones. In fact, it is likely that each subset of traits may have their own fitness functions and relate to other traits in an additive or multiplicative (or other more complex) way. Whether the fitness function is additive or multiplicative and whether it has flat areas or not has an effects on fitness dynamics.

In each simulation the optimum is a specific combination of trait values, vector t_o. The initial condition is also a vector, a different one per individual, vector t_i. The initial conditions can be: homogeneous low, optimal or random (each trait in each individual taking a random value between a low value and the optimal value with uniform distribution). In the low initial conditions t_i is chosen so that for the fitness function used each trait has a fitness contribution of 0.05. That is a low value but it is still in the area of the fitness function that is not flat.

Mutation is implemented as changes, with a rate μ per locus, in the phenotypic trait value associated with a trait, $t_{j\ new} = t_{j\ old} + m$, where m is a stochastic random variable with a normal distribution in some simulations $N(0,\sigma_t)$ and a uniform distribution U in others $(-\sigma,\sigma)$. Note that this is very different from the universal pleiotropy used in Fisher's geometric (Orr 2000; Poon & Otto, 2000; Martin & Lenormand, 2006; Tenaillon *et al.*, 2007; Lourenço *et al.*, 2011). In our model every trait varies independently. This is mutations affect only one trait at a time. In models based in Fisher's geometric model each mutation is like a random n-dimensional vector in the phenotypic space and thus most mutations affect all traits at the same time. This difference makes that in our model the likelihood of mutations that decrease the distance to the optimum does no depend on n, simply each mutation has a probability of 1/2 to decrease the distance of a trait to the optimum and a probability of 1/2 of increasing that as long as the trait is far from its optimum (the actual changes in fitness that these mutations produce depends on the fitness functions and on the distance of a trait and the rest of the traits in an individual to their optimum). We chose to consider only the simple case in which traits vary independently because again we want to study only the conditions that are most ideal for the action of natural selection.

Our model does not include genes as such; it simply considers the probability of heritable change per trait and considers only traits that vary independently (but see later). We use this simplification because our aim is to understand the long-term evolution of phenotypes and estimate the limits to phenotypic evolution in ideal conditions (e.g. simple relationship between genotype and phenotype, trait independence).

The mutation rates used in the model are much greater than nucleotidic mutation rates found in nature (Lynch, 2010) (and to the mutation rates used in multi-locus models). This is because we

actually implement mutation rates per trait (not per gene or nucleotide). In fact, mutation in our model can be understood as the chance of having an heritable change in a trait irrespective of whether, in nature, the trait is determined by a single locus, several locus or by a complex network of genetic interaction between polymorphic loci. In biological systems, since most traits are controlled by many genes, variation in a trait should appear whenever one of its many controlling genes mutates. Thus, the mutation rates used in the model can be seen as the sum of the probabilities of mutation of all underlying genes for a given trait. Our model then assumes that, on average, all traits mutate at the same rate and this mutation rate is roughly constant over time. This is again an idealization allowed because we are interested in the limit case in which everything is most ideal for selection to act.

The coefficients s_j are chosen to be all equal to one. In this sense all traits can equally contribute to individual fitness. This value represents the increase in individual fitness arising per unity of increase in a trait contribution to individual fitness, f_{ij}. For the linear fitness function this is directly equivalent to the selection differential. We chose this coefficient to be rather large in order to represent an ideal condition for the action of natural selection (although similar results are found for other values). In a similar way d_{max} is analogous to the inverse of selection gradient (b) (Lande & Arnold, 1983). We use a value of d_{max}=10 that corresponds to a selection gradient of 0.1. This is again a very large value, compared to estimated natural values (Kingsolver *et al.,* 2012), chosen to study conditions ideal for natural selection (notice that since the relationship between genotype and phenotype is very simple there are no local optima).

Natural selection and drift are implemented by choosing, at random, each individual in each new generation from the individuals in the previous generation. The probability of being chosen is equal to the relative fitness. Non-overlapping discrete generations are used.

All simulations were run for 10^7 generations and then we manually assessed whether fitness had saturated with time. If that was not the case the simulation was continued and manually followed until saturation. Most simulations never attain the optimum but in all cases the population maximal fitness reaches a steady state value that we call the maximal equilibrium fitness $\hat{W}$.

In order to assess the effect of recombination on $\hat{W}$ we performed a set of simulations in which individuals of the population paired up at random. Our simulations, as explained, do not make any assumptions on the genetics underlying trait variation. In order to apply recombination, however, something has to be assumed about the underlying genetics. Thus, in the simulations with recombination, and only in those, we assume that each trait is determined by a single unique locus and all the loci are arranged in a single chromosome.

More precisely, the rate of recombination R establishes the probability of a chromosomic

crossover to take place in a locus. All the loci between two crossover points are interchanged (despite all the traits are equivalent, in the recombination simulations they are kept ordered during simulations according to their index *j*, making possible to apply this recombination algorithm). When R=0.5 , half of the loci are interchanged, which represent the maximum effective recombination.

We also performed as set of simulations, *evolvable n simulations*, in which some mutations, at a rate $\mu_n=10^{-4}$ and 10^{-3} per loci, increase or decrease n, by +1 or -1 per mutation. When an n-increasing mutation occurs, the phenotypic value of the new trait is equal to the phenotypic value for this trait at the beginning of the simulations. This way individuals with different n compete in the population. In these simulations, mutation rates μ_n higher than μ are not considered because it is assumed that mutations leading to new phenotypic features (traits) are far rarer than those producing variation in existing traits (Goldschmidt, 1940; Kopp & Hermisson, 2009; Good *et al.*, 2012). How new traits arise in the phenotype by mutation is not very well understood, here we simply assume that this happens from time to time with low probability. It may seem reasonable that new traits may not start being totally independent from the already existing traits in a phenotype. However, since we are studying a limit model in which everything should be ideal for natural selection to act most effectively we only consider new traits that are independent from the rest. The evolvable-n simulations are used to provide precise estimations of which is the maximal number of traits that can be maintained by natural selection for specific population sizes and mutation rates and to study how the order in which traits arise in evolution affects, for different fitness functions, the relative adaptation of each trait among all the traits in an individual phenotype.

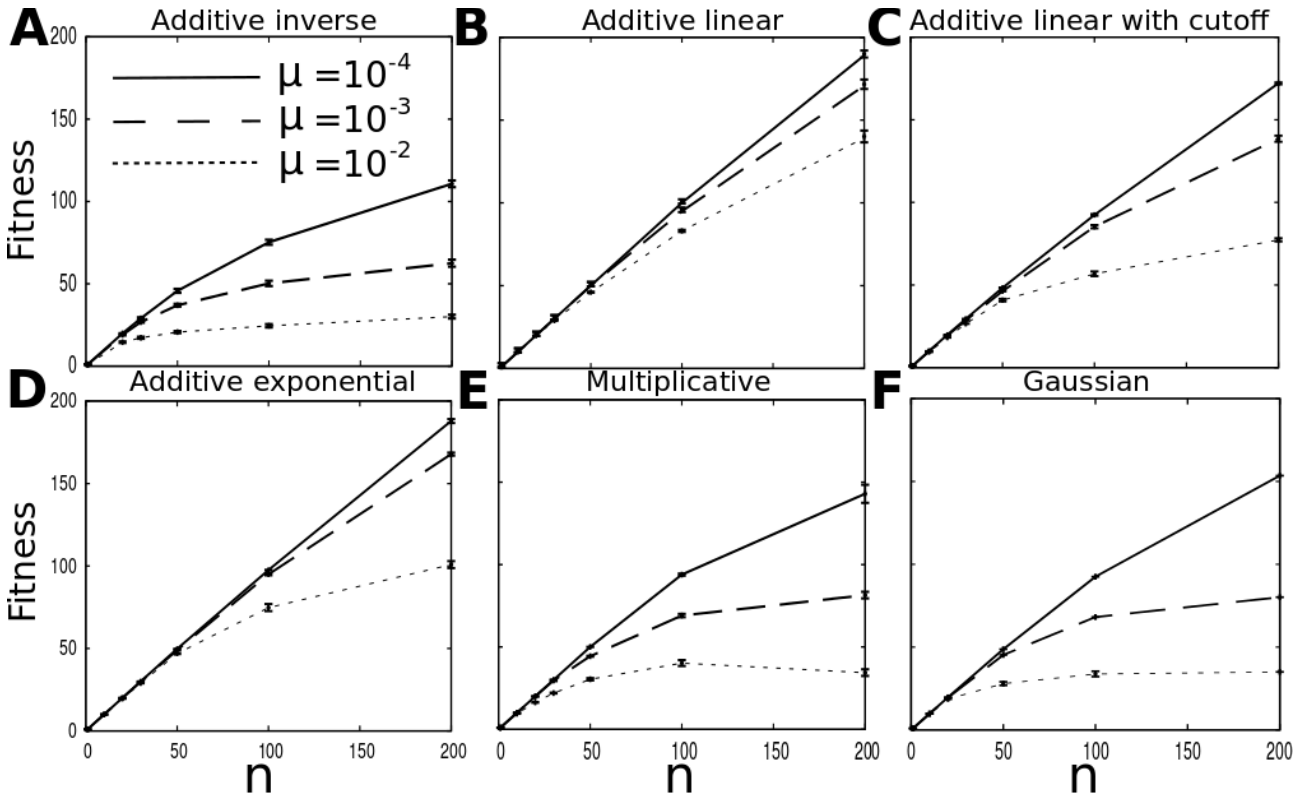


Figure 1. $\hat{W}$ is proportionally smaller for large number of traits n. Maximum absolute fitness $\hat{W}$ attained in the steady state for simulations with different n and μ for all fitness functions considered: (A) Additive inverse, (B) Additive linear, (C) Additive linear with cutoff, (D) Additive exponential, (E) Multiplicative and (F) Gaussian (fitness is in here multiplied by n to allow for comparisons with the other fitness functions). In figure S.4 the same results are plotted in respect to the optimality (that is the fitness divided by n). Each point is the average of 20 simulations, s_j values were set to 1 , P to 1000, d_{max} to 10, t_o=20.

Results:

Our simulations show that, for many fitness functions (Fig. 1), population size (Fig. 2), mutation rate (Fig. 1), distribution of mutational effects M (normal or uniform) (Fig. 3B) and initial conditions (Fig. 3A) the optimum is often not reached. This is specially the case for the non-additive fitness functions (and for the additive inverse) and for the additive fitness functions when mutation is not low (we will later suggest why high mutation rates are more realistic than these low mutation rates). For large n, many traits do not reach their optimal value, in fact, for large n the majority of traits are relatively far from their optimal value. $\hat{W}$ increases in a saturating manner with n, making the optimality $\hat{O}$ decrease with n in a non-linear manner (Fig. 1, Fig. S4).

The efficiency of natural selection decreases with n in a manner that depends on the fitness function. The rate of mutation decreases $\hat{W}$, although $\hat{W}$ is attained faster for larger μ. $\hat{W}$ increases with population size (Fig. 2). We also found that recombination has only a mild effect on $\hat{W}$ (Fig. 3C).

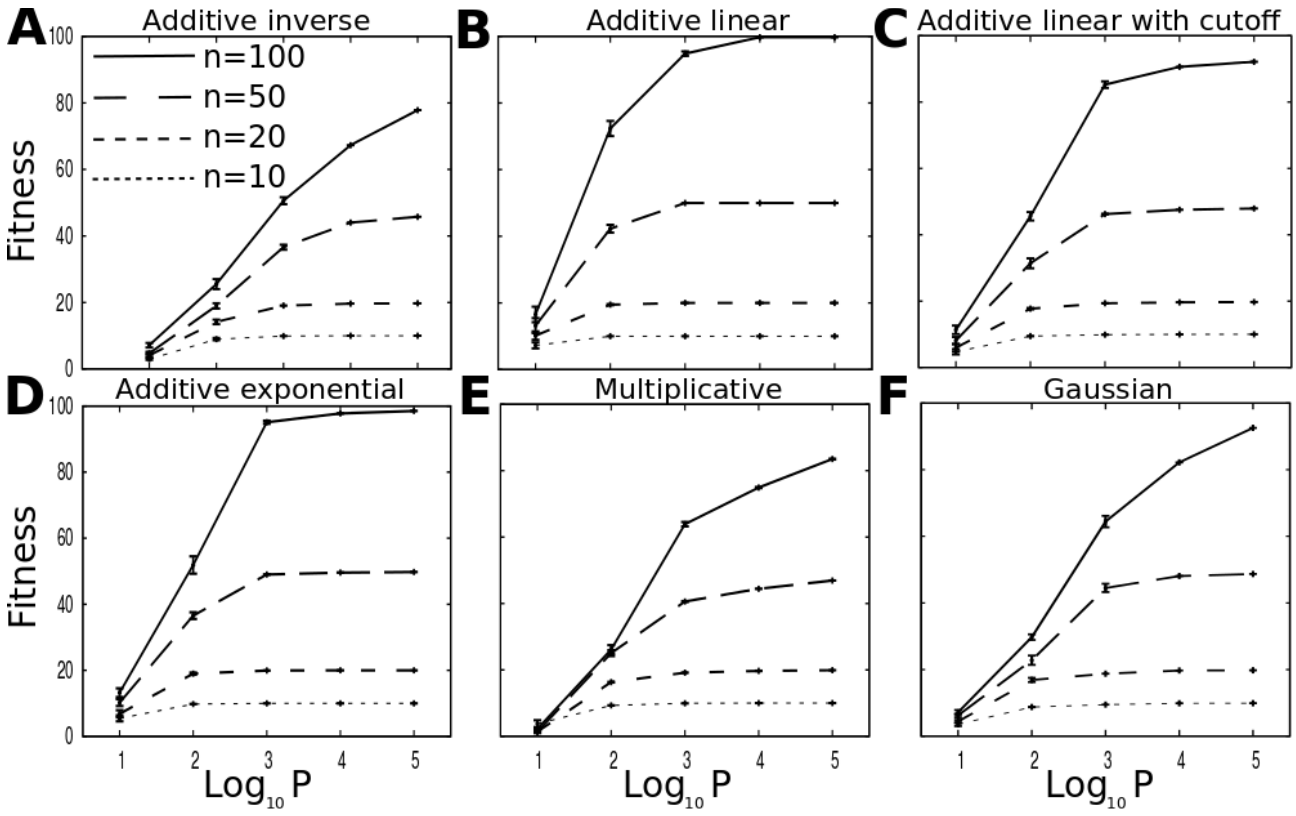


Figure 2. Plot of the maximum absolute fitness $\hat{W}$ attained for populations with different population size P and n for all fitness functions considered: (A) Additive inverse, (B) Additive linear, (C) Additive linear with cutoff, (D) Additive exponential, (E) Multiplicative and (F) Gaussian. To better appreciate which populations reach the optimum we have added figure S.5 where the optimality is shown. Each point is the average of 20 simulations, s_j values were set to 1, d_{max} to 10, μ=10^{-3}, t_o=20.

These results imply that complex phenotypes, defined in here on the bases of the number of traits, n, have lower levels of optimality. This first part of our results shows that adaptation is not only slower for large n, as the work cited in the introduction suggests, but that this is so to the extent that, even when conditions are ideal for natural selection, the optimal phenotype is often not reached and that an important proportion of phenotypic traits are largely sub-optimal. The distribution of optimalities per trait (Fig. 4) is, for all fitness functions except for the additive linear, largely skewed with some traits being optimal, or nearly so, and many traits being not adapted at all (just as if they would be neutral).

In addition, for the non-linear fitness functions, the most adapted traits tend to be also the ones that start to adapt earlier in evolution. This trend, however, fades if simulations are run for many generations (Fig. 5A).

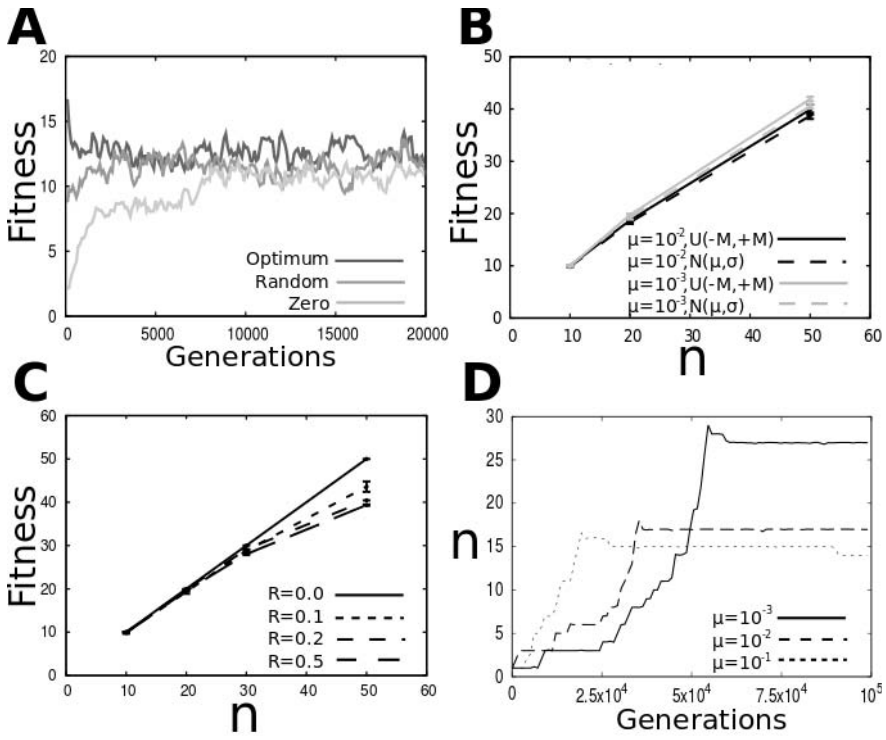


Figure 3. (A) $\hat{W}$ is independent of the initial conditions Average individual fitness over generations in example simulations (additive inverse fitness function). Three different initial conditions are depicted in each plot: homogeneous low, optimal or random (t_i coming from a uniform distribution between low and $t_{o\,j}$). μ=10^{-2}, s_j values were set to 1 and d_{max} =10, mutational distribution $N(0,\sigma_t)$ with σ_t=1. (B) Effect of the kind of mutational distribution on $\hat{W}$. Maximum absolute fitness $\hat{W}$ attained in the steady state for populations in which the distribution of new mutants follows a normal $N(0,\sigma_t)$ (dashed lines) or uniform distribution $U(-M,M)$ (solid lines). Two mutation rates are considered μ=10^{-2} (black lines) and μ=10^{-3} (gray lines). In this plot, populations with uniform distribution have slightly higher $\hat{W}$. This can be understood by considering that the real variance of a $U(-M,M)$ distribution is $(2M)^2/12$, which is smaller than the variance of the $N(0,\sigma_t)$ distribution (σ_t by definition) for all $M\leq 3$. In this case M=1. The linear fitness function was used, P=100, s_j=1 and d_{max}=10. (C) Genetic recombination does not increase $\hat{W}$. Maximum absolute fitness $\hat{W}$ attained in the steady state for simulations with different n and different recombination rates R. Note that the maximum effective R is 0.5 (see text). Each point is the average of 20 simulations. Additive linear fitness function was used, t_o=20, s_j =1, P=1000, μ= 10^{-2} and d_{max} =10. (D) In evolvable n simulations, evolution leads to a limited increase in n. That n is larger for low mutation rates. Each line is the average of 10 simulations for a given mutation rate μ. Additive linear fitness function was used, s_j values were set to 1, P to 100, d_{max} to 10, t_o=20, μ_n=10^{-4}.

In the evolvable n simulations with additive fitness functions, but not in the multiplicative one, n tends to increase over generations, n is 1 at generation 1, until a steady state is reached in which n oscillates around an equilibrium value that we call n_{max} (Fig. 3D). Since individual fitness is a sum of trait terms, having more terms allows larger fitness (no direct cost of complexity is assumed in our simple model). In the additive fitness functions totally optimal individuals get

replaced, over time, by individuals that are not optimal in all traits but that sum up to a larger absolute fitness from a larger number of non-optimal traits. Thus, increasing the phenotypic complexity *n* produces a decrease in optimality $\hat{O}$, traits are less adapted (Fig. 6). In addition, the traits that arise earlier in evolution are also the ones that are most adapted (Fig. 5C). Thus, old traits contribute a dis-proportionally large part of the overall fitness of individuals. This is not the same phenomenon than phylogenetic inertia (*sensu* Bloomberg & Garland, 2002) because in the model, all traits remain equally independent during simulations, and nor canalization nor integration between early arising traits can occur (the model is far too simple for that to be possible). Over time absolute fitness increases (Fig. 6) but its rate of growth decreases and the contribution to the individual fitness of each new trait decreases too.

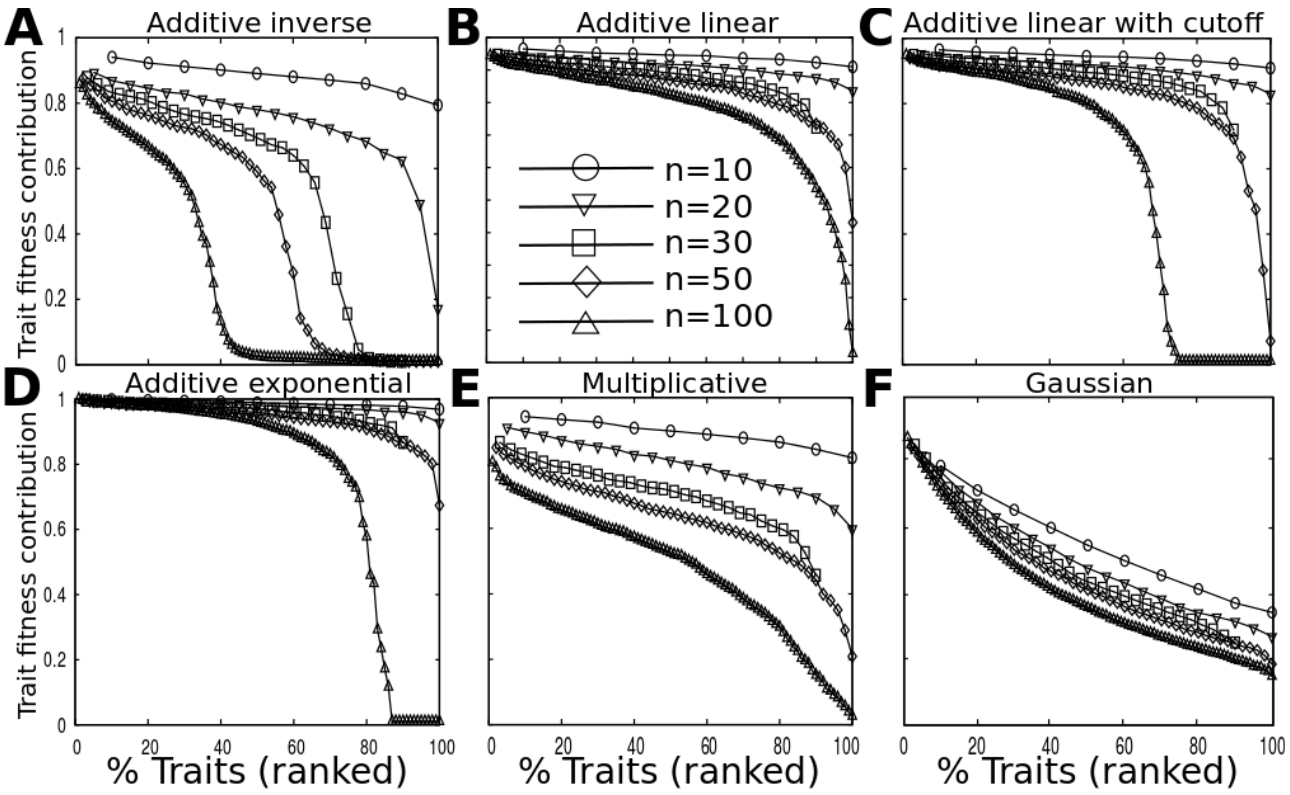


Figure 4. The proportion of traits that behave neutrally increases with *n*. Rank-ordered distribution of trait fitness in the steady state population for the different functions: (A) Additive inverse, (B) Additive linear, (C) Additive linear with cutoff, (D) Additive exponential, (E) Multiplicative and (F) Gaussian. The populational average for each ranked trait (Y-axis) was taken and plotted against that rank divided by n (X-axis). The proportion of traits that have very low fitnesses: i.e. are not adapting, increases with *n*. 20 replicates, P=1000. t_o=20, s_j=1 , μ= 10^{-2} and d_{max}=10.

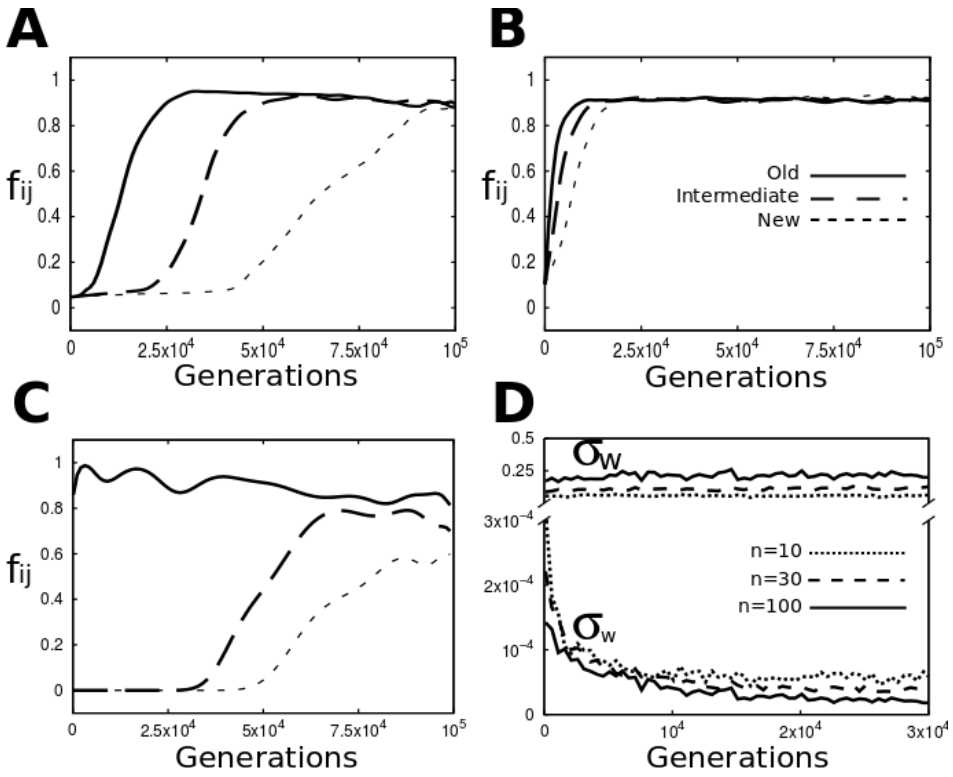


Figure 5. Traits that start to adapt earlier (this is they increase f_{ij} earlier) reach higher adaptation levels under non-linear fitness functions: (A) additive inverse; but not under linear finess functions: additive linear (B). In these plots, traits are classified according to the time they get the neighbourhood of the optimum (they reach a f_{ij}>0.5). That way the the first third of the traits which surpass this treshold are labeled as "old" traits, the second third as "intermediate" traits and the last third as "new" traits. Each line is the average of the f_{ij} of each set of traits. n=50, t_o=20, s_j =1, P=1000, μ= 10^{-3} and d_{max} =10, mutational distribution $N(0,\sigma_t)$ with σ_t=1. (C) In evolvable-n simulations, traits that arise earlier in evolution are also the ones that are most adapted. In this plot, traits are classified according to the time they appear in the population. As before, the the first third of the (n_{max}) traits that appear during simulation are labeled as "old" traits (solid line), the second third as "intermediate" traits (dashed line) and the last third as "new" traits (dotted line). In a given generation, new traits are in average, less adapted than older ones. Thus, old traits contribute a dis-proportionally large part of the overall fitness of individuals. Linear fitness function, t_o=20, s_j =1, P=100, μ= 10^{-3} ,μ_n= 10^{-4} and d_{max} =10 (D) Standard deviation of absolute and relative fitness (σ_W and σ_w) across generations an example simulation. Additive linear fitness function was used, t_o=20, s_j =1, P=1000, μ= 10^{-3} and d_{max} =10.

All these results can be understood qualitatively from the basic dynamics of drift, selection and mutation. In the Supporting Information we provide a more mathematical discussion of some of these results. We will first provide this explanations for the additive functions and later for the multiplicative one.

That absolute fitness increases with n is trivial since for large n more traits contribute to absolute fitness and thus a higher absolute fitness can be attained. That this increase should decrease

with *n*, or that, equivalently, optimality should decrease with *n*, can be understood, for the additive functions, from the fact that *n* decreases the variance of the relative fitness in the population (see Fig. 5D). This is because the relative importance, or contribution, of each trait to an individual fitness becomes necessarily smaller the more traits there are. Thus, as *n* increases, natural selection is less efficient in discerning between the different variants in a trait in proportion to the always present drift (that eliminates both adaptive and non-adaptive variants). On average, the larger *n* the more far away from its optimum each trait is and the smaller is the overall optimality. Notice that this effect is different from the one suggested by Fisher. In Fisher's model *n* decreases the chances of producing adaptive mutations and thus if the population is small it produces a drift load because *n* increases the proportion of mutations that are maladaptive (thus increasing the relative importance of drift in respect to selection). In our model that the optimum is not reached (and the distance to it that populations attain) is not due to a smaller likelihood of producing adaptive mutants as *n* increases but to the decrease in the relative fitness contribution per trait with *n*. This decreases the efficiency of effective natural selection on traits, in respect to drift, and thus precludes the optimum to be reached and increases the distance to it that is reached by a population. A similar effect is implicit in previous studies for the case of the Gaussian fitness function (Waxman and Peck, 1998; Lourenço *et al.*, 2011). These studies do not explore all the same questions we address in this article and do not explore the different dependences of optimality on *n* for other fitness functions as in here.

The mutation rate decreases $\hat{W}$ too but this is simply the well recognized mutational load (Haldane, 1937). *n* does increase this effect in a substantial way for the non-additive and additive inverse fitness functions (see Fig. S3), and in only a mild way for the other additive fitness functions. This is likely due to *n* and mutational load affecting the importance of drift at different non-linear rates (as we explain below).

That the fitness contribution of traits in the steady state is largely skewed (Fig. 4) is because mutation and drift can, by accident, bring some traits to values where the fitness function is flat or nearly flat (even if the initial conditions we use are such, as explained above, that all traits start far from the flat area of their fitness function). Note that, as we describe in the methods, it is ecologically reasonable to expect fitness functions with some flat areas. Since *n* effectively increases the importance of drift per trait the chances of getting into a flat area are larger for large *n*. In the additive inverse fitness function, for example, the contribution to individual fitness of a trait goes as $1/(1+d_{ij})$ and clearly after some distance from the optimum any decrease in distance has a negligible effect on fitness (the slope of the function approaches zero). As a result of that we have observed (results not shown) that those trait values fluctuate randomly, as a one-dimensional

random walk. These fluctuations make that some traits are effectively lost: they never, or very rarely, return to values close enough to the optimum where they can make a significant contribution to individual fitness (Fig. 4). These are the traits that show no adaptation in Fig. 4. No traits get lost in the additive linear fitness function and yet there is a skew in the f_{ij} distribution. In this case the traits with low f_{ij} are the traits that, due to genetic drift, have accumulated a large number of maladaptive mutations. Since to have really low f_{ij} a trait needs to accumulate several maladaptive mutations, the frequency of really low f_{ij} decreases non-linearly (since this involves the co-occurrence of several mutations events) as observed. Over time these traits would by chance acquire adaptive mutations and likely increase their fitness but at the same time others would accumulate deleterious mutations so that Fig. 4 shows a dynamic equilibrium situation.

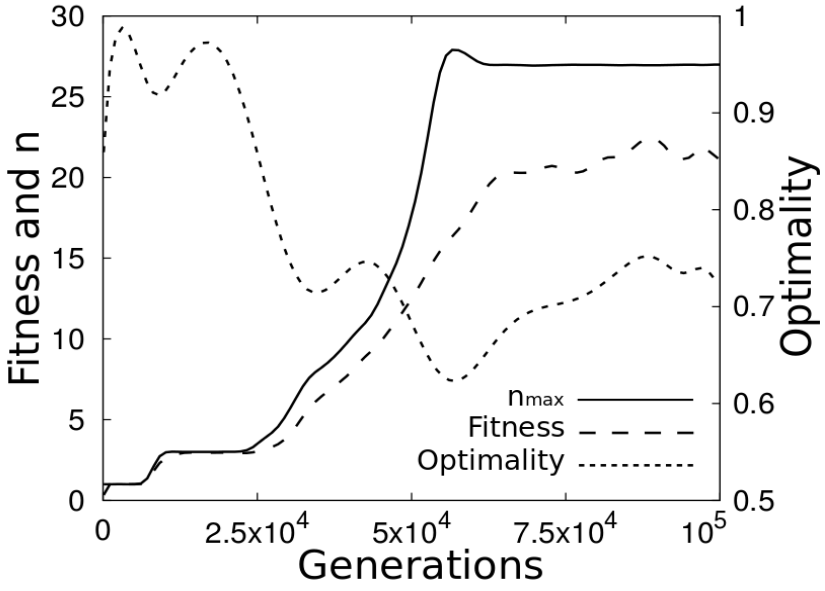


Figure 6. Optimality $\hat{O}$ over time (red line, right scale of Y-axis), showing that populations with low mutation rate μ are less optimal, despite they have higher number of traits n and fitness $\hat{W}$ (green and blue lines, left scale of Y-axis). Additive linear fitness function was used, s_j values were set to 1, P to 100, d_{max} to 10, t_o=20, μ_n=10^{-4}, mutational distribution $N(0,\sigma_t)$ with σ_t=1.

The fact that the traits that start to adapt earlier (this is they increase f_{ij} earlier) are the ones that reach higher adaptation levels can be understood from the non-linearity of most fitness functions (Fig. 5A). The traits that start to adapt earlier also adapt faster. In addition, early in each evolution simulation, many traits are far from their optimum and thus have trait values where the fitness function has a relatively low slope and do not contribute much to individual fitness. As a result early evolution occurs as if the effective *n* would be much lower since only the variation segregating in the already highly adapted traits has a significant contribution of individual fitness. Over time, however, more and more traits start to approach their optimum and the effective *n* increases and the effective contribution to individual fitness per trait decreases. This also explains why the f_{ij} of early evolving traits decreases over time after reaching a maximal value early in evolution (when only some few traits are well adapted) (Fig. 5A).

In the evolvable *n* simulations the traits that arise later in evolution start with low trait values while the older ones have had time to increase their values towards their optimum (Fig. 5C). In addition, the new traits start to evolve in a context where *n* is larger so that it is more difficult for them to increase their values towards the optimum. The increase in *n* does not affect the old traits f_{ij} if those traits are already relatively well adapted, specially for the non-linear fitness functions. This is because old traits had had the chance to reach the high f_{ij} values where the non-linear fitness functions have their larger slope. In this sense the distribution of f_{ij} values provides a historical record of which traits evolved first. Again, however, genetic drift would erase this record over time because even the most adaptive traits can, by chance, fluctuate to low fitness values and even very maladapted traits, given enough time, can by chance reach optimality. However, the more adapted a trait is the less unlikely is drift, and thus the longer time it would take to lower its value very far away from the optimum. Thus, this historical record should fade over time but it should fade first for the younger traits.

In the evolvable *n* simulations with additive fitness functions, *n* increases over generations because each increase in it adds a new term contributing to individual fitness. Thus, individuals with larger *n* sum to larger fitness values and can replace, over time, individuals with lower *n*. Ultimately, however, *n* becomes so large that the new traits arising by mutation can not effectively contribute to individual fitness and then can not adapt. At this point, since the mutations increasing or decreasing *n* are equally likely, further increases in *n* are not adaptive but neutral. As a result *n* fluctuates around the last value of *n* that provides an adaptive advantage (as seen in Fig. 3D).

n does not increase over time for the multiplicative fitness function simulations. As in the additive case, new traits start being far away from their optimum and thus have a large d_{ij} and, consequently, small fitness contributions, f_{ij}. But in the multiplicative fitness function, since this f_{ij} is

multiplied with the rest, a large reduction in individual fitness is obtained, precluding further increases in *n*.

That the same $\hat{W}$ is reached independently of the initial conditions (Fig. 3A) reflects that the limits on adaptation do not arise because of lack of available variation but because of the mutational load and the mentioned incapacity of traits to evade drift when *n* is large. Notice that, as mentioned, mutational load is always present and is not affected by *n* (so it is not informative about the questions addressed in this article). The independence of the initial conditions holds as long as the initial trait values are not in a flat area of the fitness function (as long as a trait is not very far away from the optimal value). In that latter case each trait would evolve solely by drift, as a one-dimensional random walker, until some traits reach values where the slope of the fitness function would be significantly different from zero. After that time the population would behave as in our simulations. In other word, even in this case the result would be the same if enough generations are considered.

The finding that recombination has no major effect on $\hat{W}$ (Fig. 3C) may seem to contradict classic findings in population genetics stating that recombination would greatly accelerate the rate of adaptation by breaking linkage (Weissman, 1889). In this article we are not studying the rate of adaptation (as in Mather, 1943; Felsenstein, 1965) but the degree of adaptation of organisms. In our model the rate of adaptation always stops at some point, either because the optimum is reached or because of the lack of efficiency of natural selection in counteracting drift when *n* is large and not because any limitation in the rate at which the adaptive variations are produced. This explains why recombination, that allows to produce new variants at a larger rate, does not improve adaptation in our model. In addition, our model does not consider the optimum in terms of a unique sequence of specific alleles (as for example in Kauffman & Macready, 1995) but in terms of phenotypic values that can be achieved by many different combinations of alleles (as in Kimura, 1965) and thus it is not suitable to study replacements over time at the genetic level (and the effect of recombination on those).

Similar results are found for the additive and multiplicative fitness functions but the underlying dynamics are slightly different. In the multiplicative fitness function the relative contribution of a trait on individual fitness decreases with *n*, as it occurs in the additive functions, but in addition the effect of the mutational load (Haldane, 1937) is strongly dependent of *n* (Fig. S3, Fig. S6) . In the additive functions a mutation in a trait decreases individual fitness in proportion to how many traits there are (for example if an individual is optimal it has fitness *n* and if one trait mutates to have f_{ij} zero individual fitness becomes equal to *n-1*) while in the multiplicative case the effect is not proportional to *n* (in the same example individual fitness would go from 1 to 0 if the

mutation makes the f_{ij} of any trait equal to 0). Thus, as n increases the relative effect of the mutational load becomes larger, thus precluding optimality for large n. This difference between the two kinds of functions becomes more apparent in the evolvable n simulations. Since each trait fitness contribution multiplies that of all other traits and new traits arise far away from their optimum, the overall effect of new traits is to reduce the overall fitness. As a result n does not increase over evolutionary time.

On the bases of our results we can estimate the maximal number of independent adaptive traits, n_{max}, that can be maintained by natural selection for a given combination of P and μ. This is certainly a maximal value since normally conditions are not as ideal for selection as in our model (e.g.: non-simple genotype-phenotype map). Estimations of the mutation rate per trait per generation can be obtained from the average gene length (Deutsch & Long, 1999; Lynch & Conery, 2003) multiplied by the rate of spontaneous (non-synonymous) mutation per nucleotide ($4x10^{-8}$ in *Caenorhabditis elegans and* $6x10^{-9}$ in *Mus musculus*) (Lynch, 2010) and per the average number of loci controlling a trait (2.2 in *C. elegans* and 8.7 in *M. musculus*) (Shook & Johnsson, 1999; Valdar *et al.*, 2006). This way $\mu \approx 2.8x10^{-5}$ in the case of *C. elegans* and $\mu \approx 3.3x10^{-3}$ in mouse. Effective population sizes are likely to be context specific but the existing estimations of the effective population size are $P=80000$ in *C.elegans* (Lynch, 2010) and $P_e=60000$ in mouse (Charlesworth, 2009). Since measurements of coefficients s_j are not available for most of the traits of these species, we proceed as in the rest of simulations, keeping all $s_j=1$. We acknowledge that these coefficients are probably much larger than real ones (see methods), but we are interested in an upper estimate of n_{max}. From these values we find, using the evolvable n model, a n_{max} of 36 and 67 for *C. elegans* and *M. musculus* respectively. This value comes from extrapolating the maximum n attained in simulations with evolvable n for these μ and P.

Previous studies (Martin & Lenormand, 2006; Tenaillon *et al.*, 2007) have been also used to estimate the number of independent traits that could be maintained by natural selection. These are estimated for two viruses in one case (Tenaillon *et al.*, 2007) and for a virus, a bacteria, a yeast and two animals in the other (Martin & Lenormand, 2006). Their estimations of the number of traits are roughly comparable with our estimations although they are based in quite different approaches. In one (Martin & Lenormand, 2006) they analyze the Fisher model with universal pleiotropy and a Gaussian fitness function to estimate the distribution of the fitness effect of mutations (and by using some general assumptions about the genotype-phenotype map and about selective effects). Then, they compare that distribution with the observed ones in empirical studies in mutation accumulation lines. From this they estimate which would be the effective n if that data would be produced from the dynamics observed in their model. Their model and estimations do not use specific population

sizes but assume that the effect of population size should be implicit in the estimated and observed distributions of the fitness effects of mutations.

In the other (Tenaillon et al., 2007) they also use the Fisher model with a Gaussian fitness function and, in this case, an explicitly finite populational size. On the basis of this model they estimate the likelihood of fixation of deleterious mutations for a given population size and compare it with the rate of this fixation in experiments of virus populations of known size. This way they estimate the effective n there may be in those viruses.

The differences between these two other estimations could be understood (as suggested in Lourenço *et al.*, 2011) by relaxing this universal pleiotropy assumption. The differences between our estimation and their estimation, ours being higher, is in part due to the fact that, as explained in the methods, we use fitness functions that are not Gaussian and we do not use pleiotropy

In our estimates of n_{max} in *Caenorhabditis* and *Mus* we keep μ_n an order of magnitude below μ. Probably μ_n is several orders of magnitude smaller than μ since mutation leading to new traits can be expected to be rarer that mutations changing the value of an existing trait. Because the n_{max} estimated for these organisms are upper estimates, and we do not know how large μ_n actually is, we assume that the real n_{max} can be even lower than the estimated n_{max}.

Discussion:

Our results suggest a relatively simple approach to the long controversy about the optimality of phenotypes: since all organisms have many traits it is very unlikely that they are optimal in many of their traits. Similar conclusions are reached from previous work by other researchers (Fisher, 1930; Alberch, 1982; Lande & Arnold, 1983; Salazar-Ciudad, 2007). Thus, there is an absolute limit on the number of traits than can be maintained by natural selection, and this number seems relatively small. In other words, there is a clear and low limit on the complexity that can be attained by natural selection. We think this limit is the upper limit because our model represents the most ideal conditions for the efficient action of natural selection. Our calculations suggest that less than 100 traits may be effectively well adapted for realistic ranges of P and μ. This is, organisms have certainly many more traits but only less than 100 independent traits can be made optimal by natural selection. The other traits we observe in organisms are thus non-independent (either because they co-vary as a result of the processes that generate variation (e.g.: development), because they get selected together or for the other reasons suggested by other researchers, as discussed in the introduction). This suggest that most of the organismal complexity is non-adaptive *per se* (this does not mean that selection has not acted on those, selection may have acted on those indirectly as we

explain later).

That relative lack of optimality may be specially relevant for large *n*, this is for complex phenotypes characterized by many traits. Only a small proportion of traits should be near their optimal value, especially for large *n*, and thus many traits should be maladapted or even behave neutrally. This suggests that not only organisms are not optimal but that complex organisms, characterized by more phenotypic traits are, proportionally, less optimal and many more of their traits effectively evolve by drift (this is change non-adaptively). Our calculations suggests that even for organisms with only 100 traits, one should expect that between 60% and 10% (depending on the fitness function and population size) of the traits are far from optimal (Fig.4). Our results imply an additional and a ubiquitous reason for non-optimality (beyond pleiotropy and found for all fitness functions and not just for the Gaussian one) and a clear relationship between phenotypic complexity, simply understood in here as *n*, and the proportion of traits being optimal in a phenotype.

If, as our model suggests, natural selection leads to only few traits being optimal and the rest sub-optimal or non-adapted at all, how can we then explain the current existence of complex phenotypes with many adaptive traits? The answer probably stems from the processes not included in our highly idealized model, most notably development (Alberch, 1982; Salazar-Ciudad, 2006) and traits covariation (Bock, 1959; Lande & Arnold 1983). The importance of those will be disproportionally more prevalent in complex organisms with many traits. Thus, a large part of the phenotypic complexity would be non-adaptive *per se* but would result from side effects of the processes by which the adaptive phenotypic traits are produced (e.g. developmental mechanisms) or would be selected indirectly through their co-variation with traits that are directly adaptive as suggested before Alberch, 1982; Lande & Arnold, 1983).

Another possibility is that what is being selected is a complex function of the traits values (a complex mapping between phenotype and fitness). It has been suggested that only highly degenerate mappings, many-to-one, allow for effective natural selection (Wainwright *et al.,* 2005). These are, in other words, the mappings in which many different combinations of traits values lead to similar fitness values. Similar models of natural selection with realistic genotype-phenotype maps based on our current understanding of the development also suggest that phenotypes are not optimal for most of their measurable traits (although in this case it is because that map is too complex) (Salazar-Ciudad & Marin-Riera, 2013).

Our results also suggest that complex phenotypes can arise only if fitness functions are not multiplicative or at least if this multiplication involves only a small proportion of the traits. One could expect that the fitness functions change over time since they should depend on the interaction

of the phenotype and ecology of each population. However, we cover a number of reasonable and simple functions and they all exhibit comparable results. Thus, one could expect that if the environment changes these functions over time between each other (for example to from additive lineal to additive inverse) our conclusions may still hold as long as the fitness functions are not too exotic.

In addition our results suggest that the traits accounting for a larger proportion of an individual fitness may be also the ones that arose earlier in evolution. That would imply that the recent evolution in complex organisms should involve, mostly, relatively modest increases in adaptation. Accordingly recently evolved traits would be less adapted (Fig. 5C). This suggests an scenario of the evolution of complexity in which earlier organisms were well adapted in most of their traits and complex organisms arose from those by adding progressively less and less adapted traits. Thus, the view that organisms are well adapted in most of their traits (Cain, 1964; Orzack & Sober, 1994) would hold for simple organisms, but not for complex organisms with many traits. The selective pressures on traits, either old or new, should be expected to change over time but this line of reasoning should hold unless the environment changes significantly more often for old traits versus new traits or *vice versa*.

Our results also suggest that the evolution of complexity should have been easier early on in the evolution of complex organisms, this is that it is easier to increase complexity when complexity is low (Fig. 3D). This has also been suggested previously based on simulations on how development and its genetic networks evolve (Salazar-Ciudad *et al.*, 2001). These results also imply that old traits should be less variable since they are close to their optimum and then most variation from it is likely to decrease individual fitness. This would be consistent with von Baer law that proposes that early stages of development are more conserved between the species in a group and thus likely to be more similar to the ancestral state in this group (and thus somehow older in the sense that its arose earlier in the evolution of the group) (Gould, 1977). Both von Baer and our model suggest the same, although for different reasons.

Acknowledgments:

We thank J. Jernvall, I. Salvador-Martinez, A. Matamoro, A. Barbadilla, T. Välikangas, and R. Zimm for comments. This research was funded by the Finnish Academy (WBS 1250271) and the Spanish Ministry of Science and Innovation (BFU2010-17044) to I.S.C., the Spanish Ministry of Science and Innovation (BES2011-046641) to M.B.U., and the Generalitat de Catalunya (2013FI-B00439) to M.M.R.

References:

Alberch, P. 1982. In: *Evolution and Development* (J.T. Bonner, ed) Springer, Heidelberg.

Barton, N.H. & Turelli, M. 1991. Natural and sexual selection on many loci. *Genetics* **127**: 229-55.

Bell, G. 2010. Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Phil. Trans. R. Soc. B.* **365**, 87-97.

Bloomberg, S.P. & Garland, T. 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *J. Evol. Biol*. **15**: 899-910.

Bock, W.J. 1959. Preadaptation and multiple evolutionary pathways. *Evolution*. **13**: 194-211.

Cain, A.J. 1964. The Perfection of animals. *Biol. J. Linnean Soc.* **36***: 36-63.*

Charlesworth, B. 2009. Effective population size and patterns of molecular evolution and variation. *Nat. Rev. Gen.* **10**: 195-205.

Crow,J. & Kimura, M. 1970. An introduction to population genetics theory. Harper & Row, New York.

Deutsch, M. & Long, M. 1999. Intron-exon structures of eukaryotic model organisms. *Nucleic Acid Res.* **27**: 3219-3228.

Felsenstein, J. 1965. The effect of linkage on directional selection. *Genetics.* **52**: 349.

Fisher, *R.A.* 1930. *The genetical theory of natural selection.* Clarendon, *Oxford.*

Gerrish, P.J., Colato, A., Perelson, A.S. & Sniegowski, P.D. 2007. Complete genetic linkage can subvert natural selection. *Proc. Natl. Acad. Sci. (USA)* **104**: 6266-6271.

Goldschmidt, R. 1940. The material basis of evolution (Vol. 28) Yale University Press.

Good, B.H., Rouzine, I.M., Balick, D.J., Hallatschek, O. & Desai, M.M. 2012. Distribution of fixed

mutations and the rate of adaptation in asexual populations. *Proc. Natl. Acad. Sci. (USA)* **109**: 4950-4955.

Gould, S.J. 1977. *Ontogeny and phylogeny.* Harvard University Press, Harvard.

Gould, S.J. & Lewontin, R.C. 1979. The spandrels of San-Marco and the panglossian paradigm: a critique of the adaptationist program. *Proc. R. Soc. London Ser. B.* **205**: 581-598.

Haldane, J.B.S. 1927. The mathematical theory of natural and artificial selection. *Proc. Camb. Phil. Soc.* **23**: 838-844.

Haldane, J.B.S. 1937. The effect of variation of fitness. *Am.Nat.* 337-349.

Hansen, *T.F*. & Wagner, G.P. 2001. Modeling genetic architecture: a multilinear theory of gene interaction. *Theor. Popul. Biol. **59**: 61-86.*

Hill, W.G., Barton, N.H. & Turelli, M. 2006. Prediction of effects of genetic drift on variance components under a general model of epistasis. *Theor. Popul. Biol.* **70**: 56-62.

Jones, *A*.G., Arnold, S.J. & Bürger, R. 2004. Evolution and stability of the G-matrix on a landscape with a moving optimum. *Evolution **58**: 1639-1654.*

Kauffman, S.A. & Macready, W.G. 1995. Search strategies for applied molecular evolution. *J. Theor. Biol. **173**: 427-440.*

Kimura, *M.* 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci. (USA) **54**: 731-736.*

Kimura, *M. 1983. The neutral theory of molecular evolution*. Cambridge University Press. Cambridge, UK.

Kingsolver, J.G., Diamond, S.E., Siepielski, A.M. & Carlson, S.M. 2012. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evol. Ecol.* **26**: 1101–1118

Kirkpatrick, M., Johnsson, T. & Barton, N. 2002. General models of multilocus evolution. *Genetics* **161**: 1727-1750.

Kopp, M. & Hermisson, J. 2009. The genetic basis of adaptation II: The distribution of adaptive substitutions in the moving optimum model. *Genetics* **183**: 1543-1476.

Kondrashov, A.S. & Turelli, M. 1992. Deleterious mutations, apparent stabilizing selection and the maintenance of quantitative variation. *Genetics* **132**: 603-18.

Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1210-1226.

Lourenço, J., Galtier, N., Glémin, S. Complexity, pleiotropy, and the fitness effect of mutations. *Evolution* **65-5**: 1559-1571.

Lynch, M. 2010. Evolution of the mutation rate. *Trends Genet.* **26**: 345-352.

Lynch, M. & Conery, J.S. 2003. The origins of genome complexity. *Science* **302**: 1401-1404.

Martin, G., Lenormand, T. 2006. A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* **60**:893-907.

Mather, K. 1943. Polygenic inheritance and natural selection. *Biol. Rev*. **18**: 32-64.

Maynard-Smith, J. 1978. Optimization theory in evolution. *J. A. Rev. Ecol. Syst.* **9**: 31−56.

Matuszewski, S., Hermisson J. and Kopp M. 2014. Fisher's geometric model with a moving optimum. *Evolution* 10.1111/evo.12465.

Orzack, S.H. & Sober, E. 1994. Optimality models and the test of adaptationism. *Am. Nat*. ***143: 361-380.***

Orr, H.A. 2000. Adaptation and the cost of complexity. *Evolution* **54**: 13-20.

Paley, W. 1802. *Natural theology or, evidence of the existence and attributes of the deity*. London.

Parker, G.A. & Maynard-Smith, J. 1990. Optimality theory in evolutionary biology. *Nature* **348**: 27-33.

Poon, A., Otto, S.P. 2000. Compensating for our load of mutations: freezing the meltdown of small populations. *Evolution* **54**:1467-79.

Salazar-Ciudad, I., Newman, S.A. & Solé, R.V. 2001. Phenotypic and dynamical transitions in model genetic networks I. Emergence of patterns and genotype-phenotype relationships. *Evol. Dev.* **3**: 84-94.

Salazar-Ciudad, *I. 2006.* Developmental constraints vs. variational properties: How pattern formation can help to understand evolution and development. *J. Exp. Zool. B Mol. Dev. Evol*. ***306****: 107-125.*

Salazar-Ciudad, *I. 2*007. On the origins of morphological variation, canalization, robustness, and evolvability. *Integr. Comp. Biol.* **47***:390-400*

Salazar-Ciudad, I. & Jernvall, J. 2010. A computational model of teeth and the developmental origins of morphological variation. *Nature **464**: 583-586.*

Salazar-Ciudad, I. & Marín-Riera, M. 2013. Adaptive dynamics under development-based genotype-phenotype maps. *Nature* **497**: 361-364.

Shook, D.R. & Johnson, T.E. 1999. Quantitative trait loci affecting survival and fertility-related traits in Caenorhabditis elegans show genotype-environment interactions, pleiotropy and epistasis. *Genetics* **153**: 1233-1243.

Shoval, O., Sheftel, H., Shinar, G., Hart, Y., Ramote, O., Mayo, A. *et al*. *2012. Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. Science **336**: 1157-1160.*

Solé, R.V., Salazar-Ciudad, I. & García-Fernández, J. 2002. Common pattern formation, modularity and phase transitions in a gene network model of morphogenesis. *Physica A.* **305**: 640-654.

Tenaillon O., Silander, O.K., Uzan, J.P., Chao, L. 2007. Quantifying organismal complexity using a population genetic approach. *PLoS One*. **2**:e217.

Valdar, W., Solberg, L. C., Gauguier, D., Burnett, S., Klenerman, P., Cookson, W. O. *et al*. 2006. Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat. Gen*. **38**: 879-887.

Vercken, E., Wellenreuther, M., Svensson, E.I. & Mauroy, B. 2012. Don't fall off the adaptation cliff: when assymetrical fitness selects for suboptimal traits. *Plos One* **7**: e34889.

Wagner, A. 1994. Evolution of gene networks by gene duplications: a mathematical model and its implications on genome organization. *Proc. Natl. Acad. Sci. (USA)* **91**: 4387-4389.

Wainwright, P.C., Alfaro, M.E., Bolnick, D.I., & Hulsey, C.D. 2005. Many-to-one mapping of form to function: a general principle in organismal design?. *Int. Comp. Biol.* **45**: 256-262.

Wake, D.B. & Larson, A. 1987. Multidimensional analysis of an evolving lineage. *Science* **238**: 42-48.

Weinreich, D.M. & Knies, J.L. 2013. Fisher's geometric model of adaptation meets the functional synthesis: data on pairwise epistasis for fitness yields insights into the shape and size of phenotype space. *Evolution* **67**: 2957-2972.

Weismann, A. 1889. The significance of sexual reproduction in the theory of natural selection. In: *Essays upon heredity and kindred biological problems* (E. B. Poulton, S. Schonland & A. E. Shipley, eds.) Clarendon Press. Oxford.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. 2007. Let the concept of trait be functional ! *Oikos*. **116**: 882-892.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. Sixth Int. Congr. Genet.* **1**: 356-366.

Zhivotovsky, L.A. & Feldman, M.W. 1992. On models of quantitative genetic-variability - a stabilizing selection-balance model. *Genetics* **130**: 947-955.

Supporting Information:

Appendix S1: Decrease of relative fitness standard deviation with n and P:

In any given generation, assuming for the moment that all t_o and s_j are equal to one and that the frequency of the different trait values is changing slowly, the absolute fitness of an individual can be described as:

$$W_i = \Phi\left[n\langle f_{ij}\rangle + \sum_{j=1}^{n} f\left(x_j\right)\right] \quad [8]$$

Where $\langle f_{ij}\rangle$ is the average f_{ij} (over the population), f is any of the functions to assign the fitness contribution of a trait based on d_{ij} and $f(x_j)$ is the difference between a trait fitness contribution in an individual and $\langle f_{ij}\rangle$. If most d_{ij} values in the population are far away from 0 we can safely assume that mutations in the traits of any individual are as likely to increase d_{ij} as to decrease it. As the traits evolve toward their optimum d_{ij} decreases and mutation becomes more and more likely to produce individuals less fit than the populational average (since trait values over t_o increase d_{ij} again). Then we can approximate the distribution of t_j values by a normal with standard deviation $\sigma = \sigma_t\sqrt{n}$ (variance $\sigma^2=\sigma_t^2$n, central limit theorem). d_{ij} has a similar distribution with a different expectation but the same σ.

After the few initial generations d_{ij} values are far from d_{max} and then we can assume that the fitness function is lineal and growing (the Φ function grows only for positive arguments) in both linear and linear positive fitness functions. We will restrict our analysis to the former function. The distribution of absolute fitness in the population can be approximated by a normal with expectation $n\langle f_{ij}\rangle$ and $\sigma = \sigma_t\sqrt{n}/\, d_{max}$. Since this distribution is symmetric around the mean, $n\langle f_{ij}\rangle$, (as long as the population is far away from the optimum) the total populational fitness is roughly $Pn\langle f_{ij}\rangle$. Then the relative fitness of each individual can be described simply as:

$$w_i = \frac{1}{Pn\langle f_{ij}\rangle}\Phi\left[n\langle f_{ij}\rangle + \sum_{j=1}^{N} f\left(x_j\right)\right] \quad [9]$$

Then w_i is a stochastic variable arising simply from the product of W_i by $(Pn\langle f_{ij}\rangle)^{-1}$. The σ of a random variable multiplied by a constant is that σ multiplied by that constant and since $Pn\langle f_{ij}\rangle$ is constant in each generation we can approximate the distribution of w_i by a normal distribution with expectation $1/P$ and σ:

$$\sigma_w = \frac{\sigma_t \frac{\sqrt{n}}{d_{max}}}{Pn\langle f_{ij}\rangle} = \frac{\frac{\sigma_t}{d_{max}}}{P\sqrt{n}\langle f_{ij}\rangle} \qquad [10]$$

Thus, the standard deviation of the relative fitness decreases non-linearly with n (fig. 5D show how σ_w and σ_W change over time and with n). The fact that this decreases also with $\langle f_{ij}\rangle$ is simply because as the optimum is approached, and so d_{ij} becomes smaller, the relative increases in fitness necessarily becomes smaller (since in our model the magnitude of trait mutations is constant over time).

Since the effectiveness of natural selection depends on the standard deviation of the relative fitness between the individuals in the population our results indicate that as n increases the population approaches the optimum more slowly. This by itself would not preclude the optimum from being reached. Mutation, however, decreases the capacity to be close to the optimum, and thus $\hat{W}$, by constantly generating individuals that are less adapted than the average. Since n decreases the standard deviation of the relative fitnesses, natural selection becomes less efficient, for large n, at filtering out these less fit individuals that arise by mutation. As a result, for large n, a larger number of non-adaptive mutations are kept in the population and, consequently, the average absolute fitness is relatively low. Since the proportion of new trait values produced by mutation is independent of n (μ is defined per trait) the increase of absolute fitness with n decreases both with n and μ, as observed in our simulations. Similar arguments can be used for the other fitness functions.

A similar argument applies to the saturation of $\hat{W}$ with P. It is well known that large populations allow larger rates of evolutionary change. This is still the case in our simulations, larger populations reach their larger $\hat{W}$ values and they do it faster. However, the increase of $\hat{W}$ with P becomes slower and slower when P is large because, as in the case of n, P reduces the standard deviation of the relative fitness (fig. S1) and thus makes natural selection less efficient in eliminating individuals that are less fit than the average. Natural selection is still more efficient for larger populations (fig. 2) simply because in larger populations there is less genetic drift.

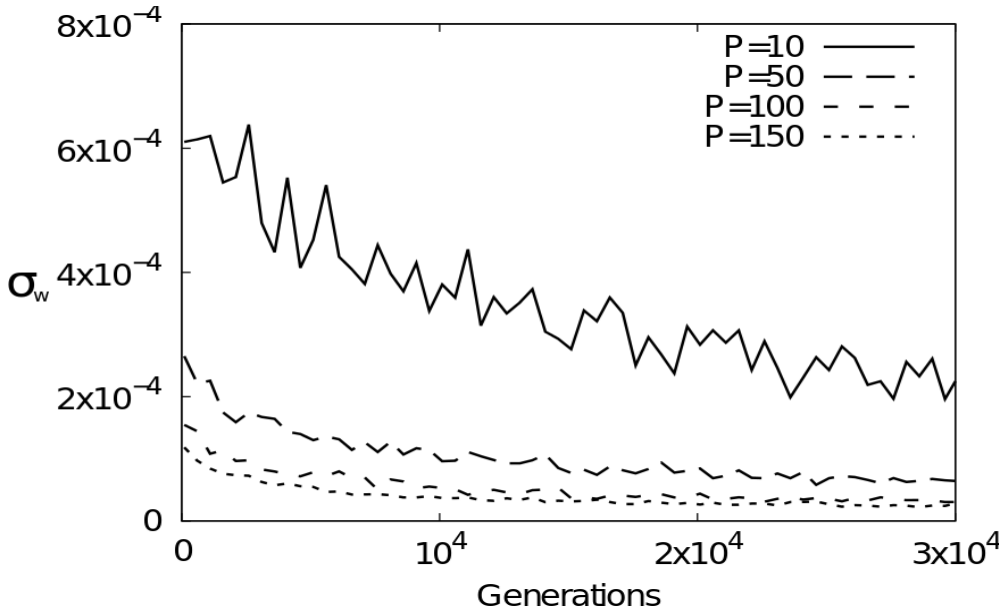


Figure S1. Population size *P* reduces the standard deviation of the relative fitness σ_w. Plot of the standard deviation of relative fitness σ_w along generations for three different population sizes *P*. Each line is the average of 10 replicates. Linear fitness function was used, s_j values were set to 1, d_{max} to 10, t_o=20, n=100, μ=10^{-3}, mutational distribution $N(0,\sigma_t)$ with σ_t=1.

Appendix S2: The effect of s_j and its standard deviation on $\hat{W}$:

If the s_j values are not all equal to one our argument is simply modified by the total absolute fitness in the population being $SPn\langle f_{ij}\rangle$ (where *S* is the sum of the s_j of all traits in all individuals). Note, however, that a larger S does not *per se* increase σ_w.

Appendix S3: Deriving $\hat{W}$ from n, μ and σ_t

In here we provide a rough approximation to $\hat{W}$ on the bases of n, μ and σ_t. Let assume a population composed of single-trait individuals with an average fitness $\langle f_i\rangle$. We consider, for simplicity and tractability, that all s_j terms are equal to 1 and that then the maximum fitness is 1. We would consider only the case in which the distribution of mutational effect is uniformly U~[-*M*,*M*] distributed and the additive linear fitness function.

In any given moment the population is placed at a certain distance to the optimum t_o. In the additive linear fitness function any increase in t_i (that is t_{i1} but since we only have a trait we simplified it as t_i) gives rise to a proportional increase in fitness (as long as t_i is not very close to t_o). In fact this increase would be m/d_{max} and have a maximum that we call $b= M/d_{max}$. As long as

$<t_i>+M<t_o$ the population is far away from the optimum and the distributions of fitness contributions are symmetric around $<f_i>$ (where $<t_i>$ is the average of the trait in the population). When $<t_i>+M>t_o$ the population is close to the optimum and the phenotypic values of some individuals surpass t_o. As a result close to the optimum the distribution of fitness contributions becomes asymmetric with more mutations decreasing fitness than mutations increasing it. In that situation we can calculate the average change in fitness due to mutation per generation by comparing the average fitness produced by the mutations over $<f_i>$ and under $<f_i>$. The mutations that decrease fitness by lowering t_i have an average fitness of $<f_i>-b/2$ (this comes from the distribution of m being uniform). For the mutations increasing t_i the situation is more complex. Some mutations, the ones increasing t_i by less than t_o-t_i (the distance to the optimum), increase the fitness while the others, the ones increasing t_i over that distance, decrease it. The average fitness of the former is:

$$<W_{pre}>=<f_i>+\frac{1-<f_i>}{2} = \frac{1+<f_i>}{2} \qquad [11]$$

This is because the fitness in the optimum is 1 and the distance to the optimum, in fitness, is $1-<f_i>$ (and thus on average that divided by 2). The mutants that pass over the optimum decrease fitness and thus produce a rebound effect. This can be calculated by considering that without that rebound effect (like if the optimum would be much more far away) there would be mutants until a distance of b from $<f_i>$. The distance past the optimum that b reaches is thus $b-(1-<f_i>)$. This is the distance between $<f_i>$ and the optimum, $1-<f_i>$, subtracted from the distance b. This is the rebound distance and thus this rebound reaches a lowest fitness of $1-(b-1+<f_i>)$ (that is $2-b-<f_i>$). Thus, the average fitness of the mutants that pass the optimum is 1/2 of the distance between the optimum and maximal rebound:

$$<W_{past}>=1-\frac{b-(1-<f_i>)}{2} \qquad [12]$$

In order to calculate the average fitness of the mutations increasing t_i these two averages, $<W_{pre}>$ and $<W_{past}>$ need to be weighted by the proportion of mutants in each class. Since in a uniform distribution all mutants are evenly spaced between $<f_i>-b$ and $<f_i>+b$, and consequently between $<f_i>$ and $<f_i>+b$, this weight can be calculated by the longitude of each class of mutants, that is $1-<f_i>$ for the mutants over t_i and under t_o and $b-1+<f_i>$ for the mutants past t_o. This longitude needs to be divided by $2b$ to give an actual proportion and taking into account that only half of the mutations decrease t_i. Thus, the average increase in fitness produced by the mutations over t_i is:

$$<W_{up}> = \left(\frac{b-F}{2b}\right)\left(1-\frac{b-F}{2}\right)+\left(\frac{F}{2b}\right)\left(1-\frac{F}{2}\right) \quad [13]$$

Where $F=1-<f_i>$.This can be simplified to:

$$<W_{up}> = \frac{1}{2}-\frac{b}{4}+\frac{F}{2}-\frac{F^2}{2b} \quad [14]$$

Thus, the average increase in fitness by mutation in each generation is:

$$\Delta W_{\mu} = \mu\left(\frac{1}{2}-\frac{b}{4}+\frac{F}{2}-\frac{F^2}{2b}+<f_i>-\frac{b}{2}-<f_i>\right) = \frac{\mu}{2}\left(1-\frac{3b}{2}+F-\frac{F^2}{b}\right) \quad [15]$$

For simplicity we consider that as μ is very small each trait only mutates once per generation, in reality it would mutate with the frequency $\mu- \mu^2- \mu^3$ … that for small μ can be simplified to μ.

This needs to be generalized to several traits. F should be, on average over all simulations, the same for all traits and thus the change in fitness due to mutation should be nF. Then the increase in fitness ΔW_{μ} in a generation coming from mutation can be approximated as:

$$\Delta W_{\mu}=\frac{n\mu}{2}\left(1-\frac{3b}{2}+F-\frac{F^2}{b}\right) \quad [16]$$

To calculate the increase in fitness produced by natural selection in each generation we have to calculate the variance of the relative fitness. We will start again for the case $n=1$. The variance of the absolute fitness is:

$$\sigma_W = E[W^2]-E[W]^2 \quad [17]$$

The average absolute fitness after mutation is:

$$E[W]^2 = \left((1-\mu)<f_i>+\frac{\mu}{2}\left(1-\frac{3b}{2}+F-\frac{F^2}{b}\right)\right)^2 = \left((1-\mu)(1-F)+\frac{\mu}{2}\left(1-\frac{3b}{2}+F-\frac{F^2}{b}\right)\right)^2 \quad [18]$$

The other term can be calculated as:

$$E[W^2]=\int x^2 f(x)dx \quad [19]$$

In our case only $n\mu$ of the traits are not on the average. The rest are on the average $1-F$. Thus:

$$E[W^2]=\mu\int x^2 f(x)dx+(1-\mu)(1-F) \quad [20]$$

Where $f(x)$ is in this case the density function of the uniform distribution and x is each possible fitness value. To include the bouncing we split the integral in two. One going from $<f_i>-b$ to $2-b-<f_i>$ and one from $2-b-<f_i>$ to 1 (that is from $1-b-F$ and $1-b+F$). The point $2-b-<f_i>$ is the lowest absolute fitness achieved by the bouncing so after this point the each fitness value is twice as likely:

$$E\left[W^2\right]=\mu\int x^2 f(x)\,dx \;+\; \mu\int x^2 2\mathrm{f}(x)\,dx+(1-\mu)(1-F) \qquad [21]$$

The uniform distribution density functions is 1/(2*b*) since all points are equally likely (just their fitness gets transformed by the bouncing):

$$E\left[W^2\right]=\mu\int x^2\frac{1}{2\mathrm{b}}\,dx \;+\; \mu\int x^2\frac{2}{2\mathrm{b}}\,dx+(1-\mu)(1-F) \qquad [22]$$

Since the density function does not depend on the fitness it can be taken outside of the integral and we can solve the simple integral:

$$E\left[W^2\right]=\frac{\mu}{2\mathrm{b}}\left[\frac{x^3}{3}\right]_{1-b-F}^{1-b+F} \;+\; \frac{2\mu}{2\mathrm{b}}\left[\frac{x^3}{3}\right]_{1-b+F}^{1}+(1-\mu)(1-F) \qquad [23]$$

$$E\left[W^2\right]=\frac{\mu}{6\mathrm{b}}\left((1-b+F)^3-(1-b-F)^3\right) \;+\; \frac{\mu}{3\mathrm{b}}\left(1-(1-b+F)^3\right)+(1-\mu)(1-F) \qquad [24]$$

Then the variance of the absolute fitness is:

$$\sigma_W=\frac{\mu}{6\mathrm{b}}\left(-(1-b+F)^3-(1-b-F)^3+2\right)+(1-\mu)(1-F)-\left((1-\mu)(1-F)+\frac{\mu}{2}\left(1-\frac{3\mathrm{b}}{2}+F-\frac{F^2}{b}\right)\right)^2 \qquad [25]$$

Expanding the previous expression:

$$\sigma_W=-\frac{F^4\mu^2}{4\mathrm{b}^2}+\frac{3\mathrm{F}^3\mu^2}{2\mathrm{b}}-\frac{F^3\mu}{b}-\frac{F^2\mu^2}{2\mathrm{b}}-3\mathrm{F}^2\mu^2+4\mathrm{F}^2\mu-F^2+\frac{9\mathrm{b}\,F\mu^2}{4}-\frac{3\mathrm{bF}\,\mu}{2}+\frac{3\mathrm{F}\,\mu^2}{2}$$

$$-3\mathrm{F}\,\mu+F-\frac{9\mathrm{b}^2\mu^2}{16}-\frac{3\mathrm{b}\,\mu^2}{4}+\frac{b^2\mu}{3}+\frac{b\mu}{2}-\frac{\mu^2}{4}+\mu \qquad [26]$$

For several traits the this variance should be, according to the central limit theorem, $\sigma_{Wn}=\sigma_W\sqrt{n}$. As before the variance of the relative fitness should then be:

$$\Delta W_s=\sigma_{wn}=\frac{\sigma_{Wn}}{P(1-F)\sqrt{n}} \qquad [27]$$

This is an approximation in which we consider a force, mutation, constantly taking away the population from the optimum and a force, natural selection, constantly bringing the population towards the optimum. When $\hat{W}$ is reached it means that the population is in a steady-state equilibrium in which mutation is compensated by natural selection. Thus:

$$\Delta W=\Delta W_\mu+\Delta W_s=0 \qquad [28]$$

We put together the contribution coming from natural selection and mutation:

$$\frac{n\mu}{2}\left(1-\frac{3\mathrm{b}}{2}+F-\frac{F^2}{b}\right)+\frac{\sigma_{Wn}}{P(1-F)\sqrt{n}}=0 \qquad [29]$$

Regrouping terms we obtain:

$$(\xi-\xi F)\frac{\mu}{2}\left(1-\frac{3\mathrm{b}}{2}+F-\frac{F^2}{b}\right)+\sigma_{Wn}=0 \qquad [30]$$

Where

$$\xi = Pn^{3/2} \quad [31]$$

Regrouping terms we obtain:

$$\frac{\xi\mu}{b}F^3 - \frac{\xi\mu(b+1)}{b}F^2 + \frac{3\xi b\mu}{2}F + \frac{\xi\mu(2-3b)}{2} + 2\sigma_{Wn} = 0 \quad [32]$$

And by substituing the laste term by Eq. [26], we obtain:

$$\frac{\xi\mu}{b}F^3 - \frac{\xi\mu(b+1)}{b}F^2 + \frac{3\xi b\mu}{2}F + \frac{\xi\mu(2-3b)}{2} - \frac{F^4\mu^2}{2b^2} + \frac{3F^3\mu^2}{b} - \frac{2F^3\mu}{b} - \frac{F^2\mu^2}{b} - 6F^2\mu^2$$

$$+8F^2\mu - 2F^2 + \frac{9bF\mu^2}{2} - 3bF\mu + 3F\mu^2 - 6F\mu + 2F - \frac{9b^2\mu^2}{8} - \frac{3b\mu^2}{2} + \frac{2b^2\mu}{3} + b\mu - \frac{\mu^2}{2} + 2\mu = 0 \quad [33]$$

We know that F is always smaller than 1 and that when the population is approaching the optimum F is smaller than b and much smaller than 1. If we discard any F term of an order larger than 1 we obtain:

$$\frac{3\xi b\mu}{2}F + \frac{\xi\mu(2-3b)}{2} + \frac{9bF\mu^2}{2} - 3bF\mu + 3F\mu^2 - 6F\mu$$

$$+2F - \frac{9b^2\mu^2}{8} - \frac{3b\mu^2}{2} + \frac{2b^2\mu}{3} + b\mu - \frac{\mu^2}{2} + 2\mu = 0 \quad [34]$$

Since μ is always very small this equation can be further approximated to:

$$\frac{3\xi b\mu}{2}F + \frac{\xi\mu(2-3b)}{2} - 3bF\mu - 6F\mu + 2F + \frac{2b^2\mu}{3} + b\mu + 2\mu = 0 \quad [35]$$

So we can isolate F:

$$F = \frac{-\frac{\xi\mu(2-3b)}{2} - \frac{2\mu b^2}{3} - b\mu - 2\mu}{\frac{3\xi b\mu}{2} - 3b\mu - 6\mu + 2} = \frac{-\mu\left(\frac{\xi(2-3b)}{2} + \frac{2b^2}{3} + b + 2\right)}{\mu\left(\frac{3\xi b}{2} - 3b - 6\right) + 2} \quad [36]$$

By substituting the F term by $1-\langle f_i \rangle$ and the ξ term by $Pn^{3/2}$, we obtain:

$$\langle f_i \rangle = 1 + \frac{\mu\left(\frac{Pn^{3/2}(2-3b)}{2} + \frac{2b^2}{3} + b + 2\right)}{\mu\left(\frac{3Pn^{3/2}b}{2} - 3b - 6\right) + 2} \quad [37]$$

We have explored the behavior of this equation for all the parameters involved (population size P, number of traits n, $b=M/d_{max}$ and mutation rate μ). We have found qualitative agreement between the analaytic solutions for $\langle f_i \rangle$ and the data of the simulations, within the range of values used in the model (fig. S2). Note that, since $\langle f_i \rangle$ was defined in our analytical approach as the average fitness contribution of a trait (see above), it has to be compared with Optimality $\hat{O}$ (that, is,

the mean absolute fitness divided by the number of traits n). $\langle f_i \rangle$ is only 1 when n=1, and decreases in a non-linear manner as the number of traits n, the mutation rate μ or b increase. However, this equation does not predict the effect population size P of on $\langle f_i \rangle$. This disagreement comes from the fact that we have not considered genetic drift in our analytical calculations.

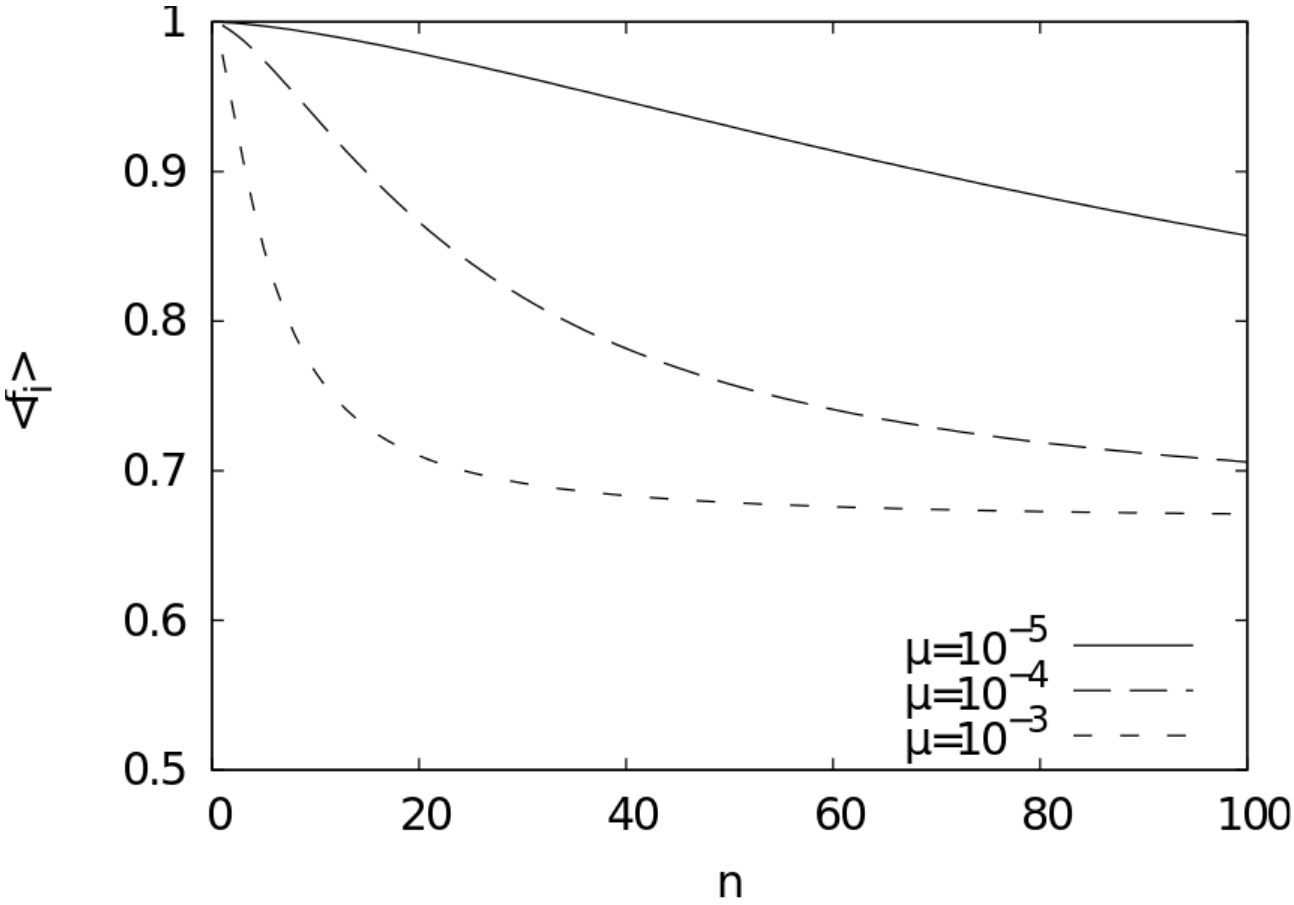


Figure S2. Analytical results: $\langle f_i \rangle$ decreases as the number of trait n increases. By solving equation 37 (see Appendix A) for different mutation rates μ and number of traits n (within the range of values used in the model), the behavior of $\langle f_i \rangle$ is in qualitative agreement with numerical simulations. $P=100$ and $b(=M/d_{max})=1$. Further agreement could be reached by considering genetic drift, but we have not been able to do it.

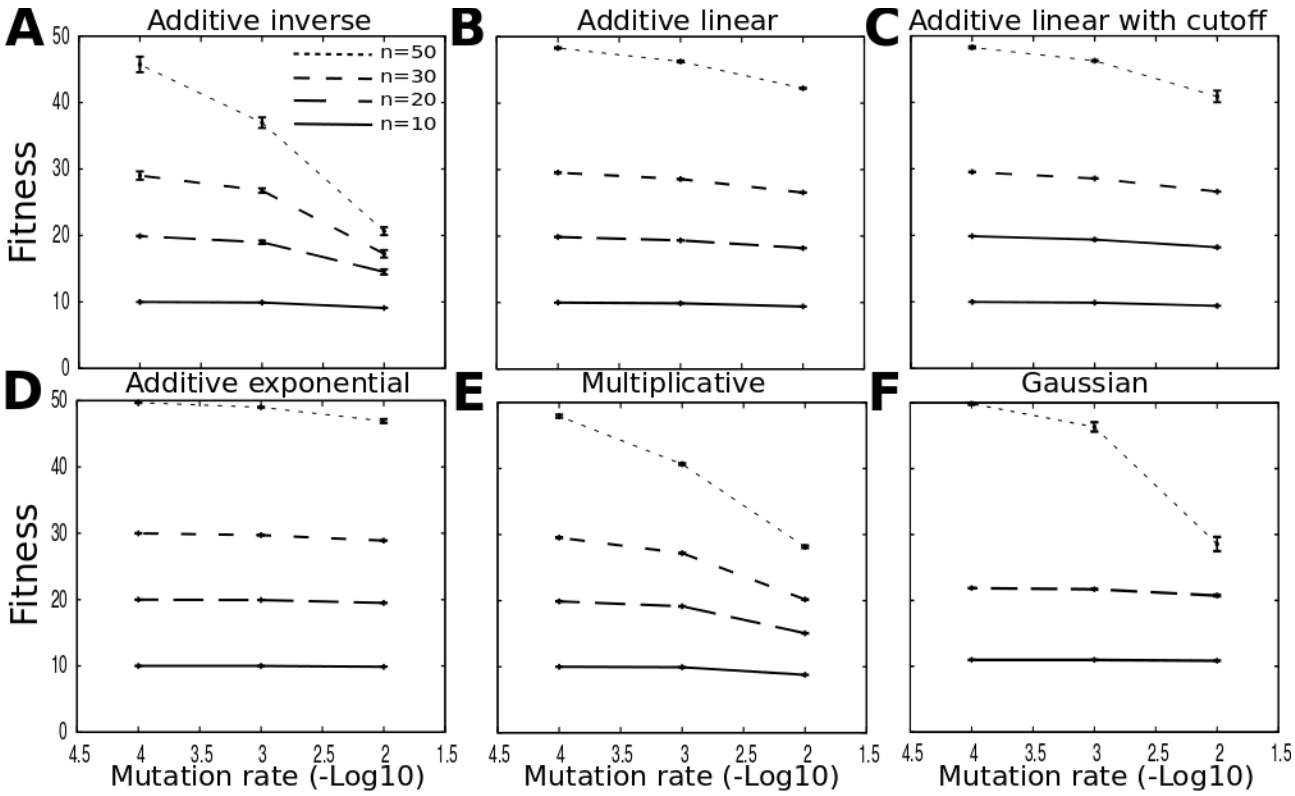


Figure S3. *n* does not substantially affect mutational load for the additive fitness functions. High mutation rates μ only decrease the maximum absolute fitness $\hat{W}$ attained when *n* is large. For moderate μ and number of traits *n*, as well as for linear fitness functions, the negative effect of μ in the $\hat{W}$ attained for a given *n* vanishes. Plots show the maximum absolute fitness $\hat{W}$ attained in the steady state for simulations with different μ and *n* for all fitness functions considered: (A) Additive inverse, (B) Additive linear, (C) Additive linear with cutoff, (D) Additive exponential, (E) Multiplicative and (F) Gaussian. Each point is the average of 20 simulations, s_j values were set to 1 , *P* to 1000, d_{max} to 10, t_o=20, mutational distribution $N(0,\sigma_t)$ with σ_t=1.

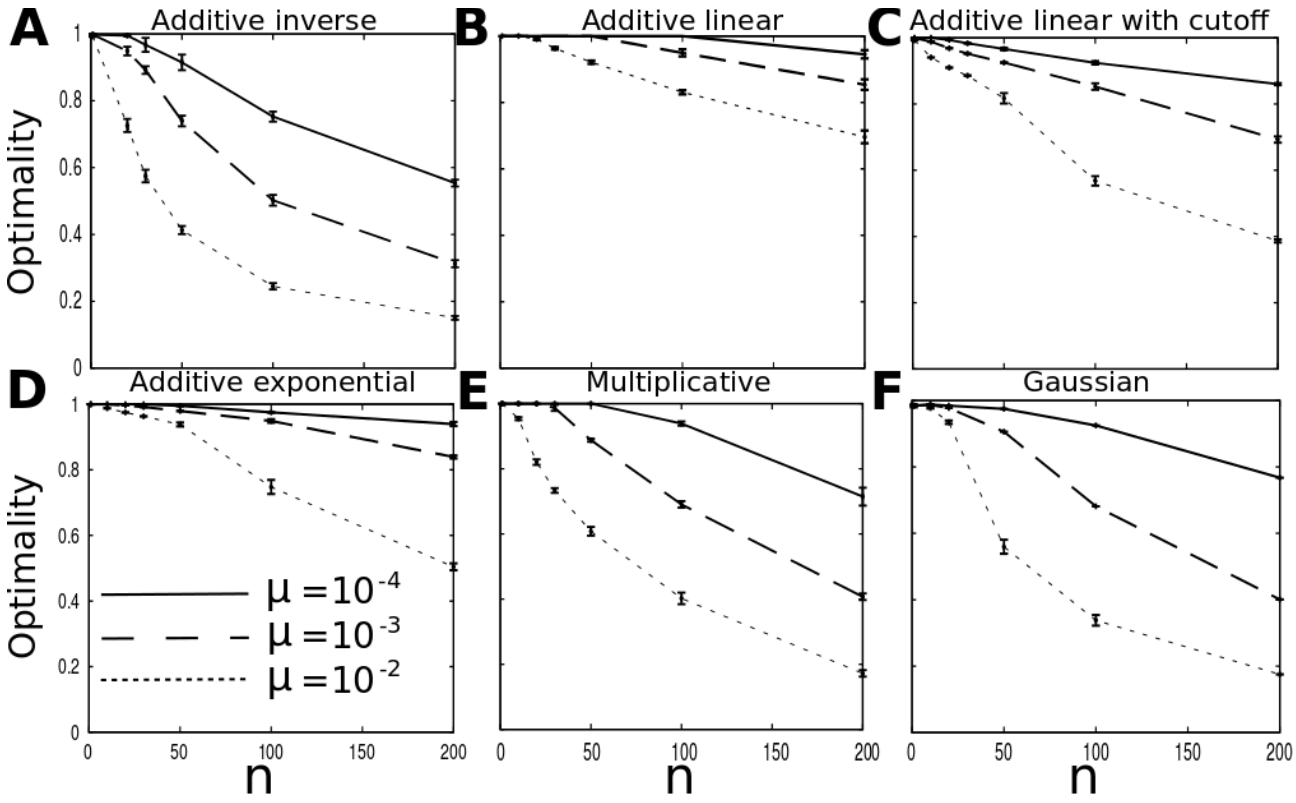


Figure S4. Optimality is smaller for large number of traits *n*. Maximum optimality attained in the steady state for simulations with different *n* and μ for all fitness functions considered: (A) Additive inverse, (B) Additive linear, (C) Additive linear with cutoff, (D) Additive exponential, (E) Multiplicative and (F) Gaussian. This figure shows the same results than figure 1, but in here Y axis corresponds to Optimality (that is the fitness divided by *n*) instead of fitness $\hat{W}$. Each point is the average of 20 simulations, s_j values were set to 1 , *P* to 1000, d_{max} to 10, t_o=20, mutational distribution $N(0,\sigma_t)$ with σ_t=1.

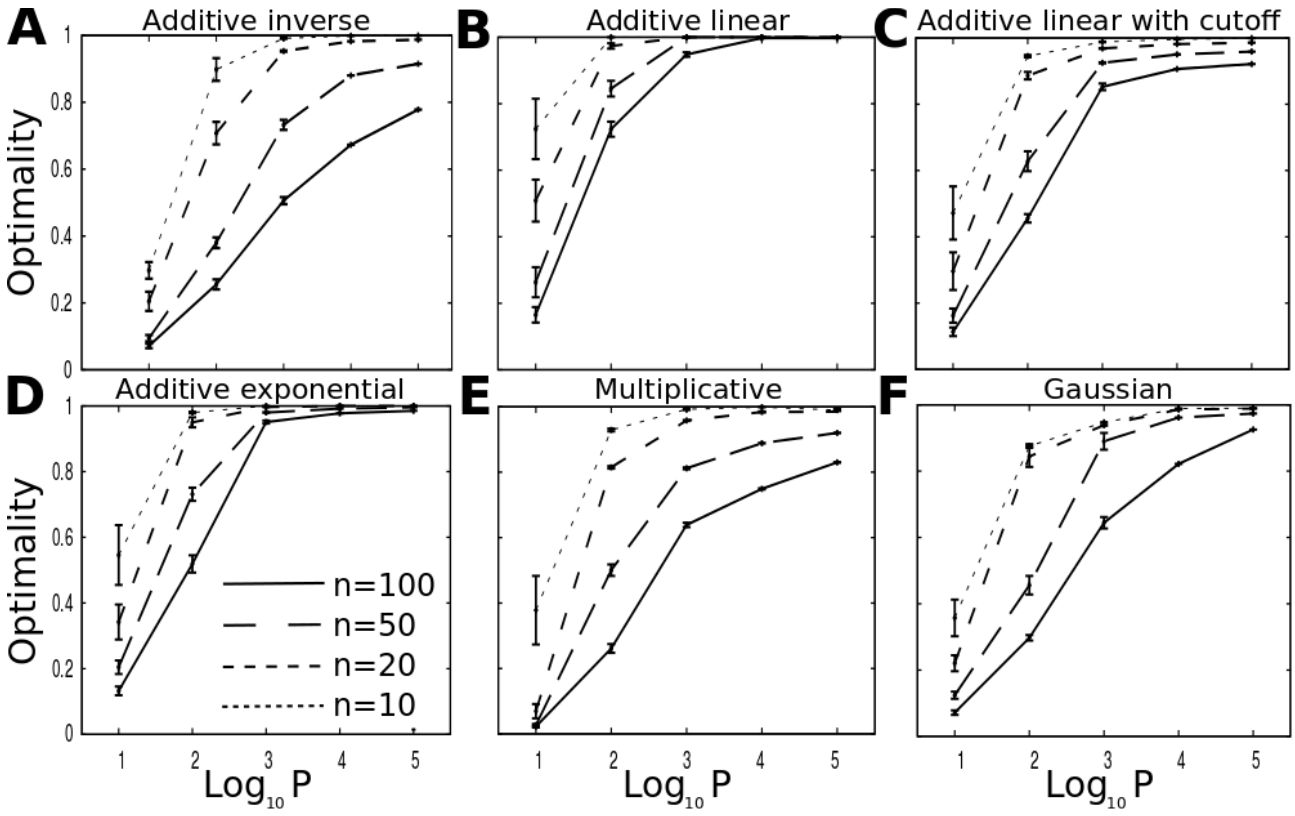


Figure S5. Plot of the maximum optimality attained for populations with different population size *P* and *n* for all fitness functions considered: (A) Additive inverse, (B) Additive linear, (C) Additive linear with cutoff, (D) Additive exponential, (E) Multiplicative and (F) Gaussian. This figure shows the same results than figure 2, but in here Y axis corresponds to Optimality (that is the fitness divided by *n*) instead of fitness $\hat{W}$. Each point is the average of 20 simulations, s_j values were set to 1, d_{max} to 10, μ=10^{-3}, t_o=20, mutational distribution $N(0,\sigma_i)$ with σ_i=1.

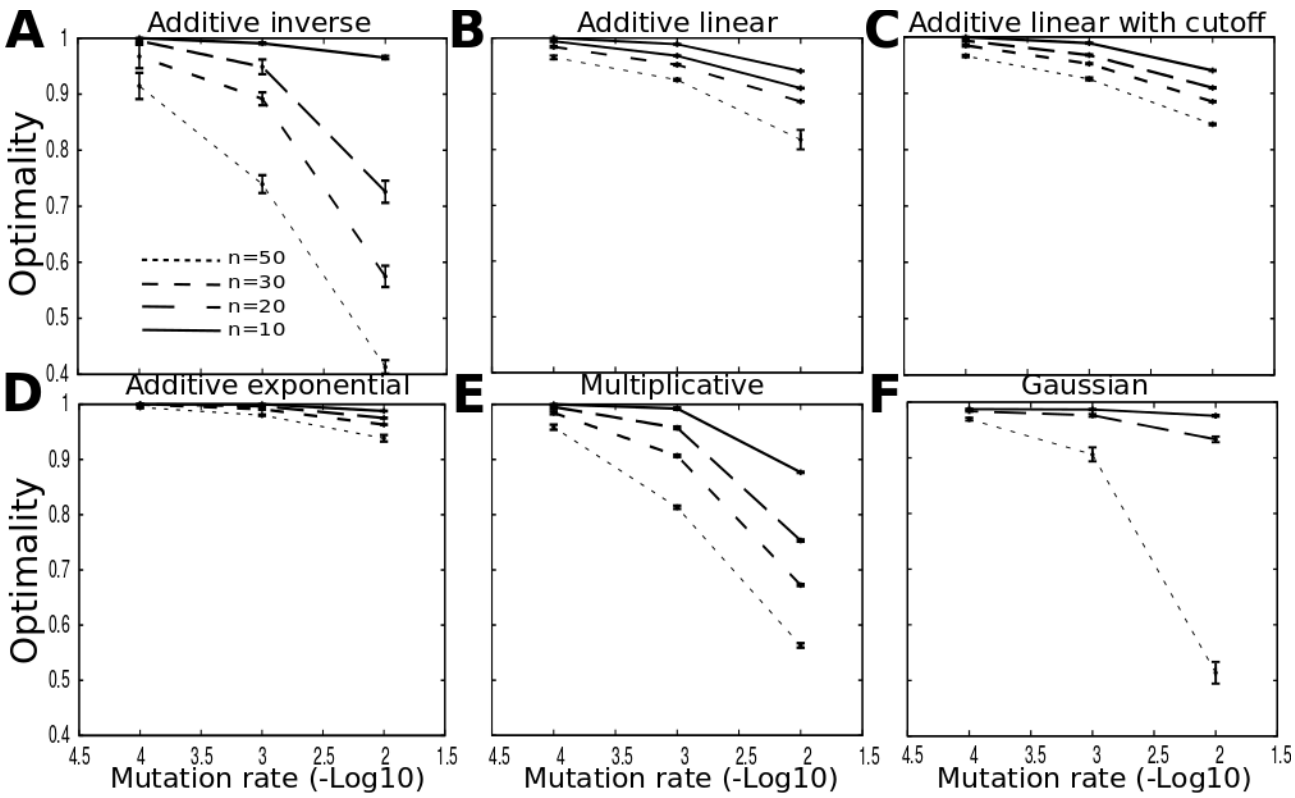


Figure S6. *n* does not substantially affect mutational load for the additive fitness functions. High mutation rates μ only decrease the Optimality when *n* is large. For moderate μ and number of traits *n*, as well as for linear fitness functions, the negative effect of μ in the $\hat{W}$ attained for a given *n* vanishes. Plots show the maximum Optimality attained in the steady state for simulations with different μ and *n* for all fitness functions considered: (A) Additive inverse, (B) Additive linear, (C) Additive linear with cutoff, (D) Additive exponential, (E) Multiplicative and (F) Gaussian. This figure shows the same results than figure S3, but in here Y axis corresponds to Optimality (that is the fitness divided by *n*) instead of fitness $\hat{W}$. Each point is the average of 20 simulations, s_j values were set to 1 , *P* to 1000, d_{max} to 10, t_o=20, mutational distribution $N(0,\sigma_t)$ with σ_t=1.

This Thesis was deposited for acceptation in the Universitat Autònoma de Barcelona on the 27th of September, 2016, and it will be defended in the same institution on the 15th of November 2016.