The principle of continuous biological information flow as the fundamental foundation for the biological sciences. Implications for ageing research.

Short-title: Life as a continuous flow of information.

Xavi Marsellach¹

¹ Independent Researcher, (L' Hospitalet de Llobregat, Barcelona, Catalan Republic).

* Correspondence to: xavi.marsellach@gmail.com

Abstract

The current state of biological knowledge contains an unresolved paradox: life as a continuity in the face of the phenomenon of ageing. In this manuscript I propose a theoretical framework that offers a solution for this apparent contradiction. The framework proposed is based on a rethinking of what ageing is at a molecular level, as well as on a rethinking of the mechanisms in charge of the flow of information from one generation to the following ones. I propose an information-based conception of ageing instead of the widely accepted damage-based conception of ageing and propose a full recovery of the chromosome theory of inheritance to describe the intergenerational flow of information. Altogether the proposed framework allows a precise and unique definition of what life is: a continuous flow of biological information. The proposed framework also implies that ageing is merely a consequence of the way in which epigenetically-coded phenotypic characteristics are passed from one generation to the next ones.

1. Life as a continuity vs the intrinsic discontinuity that ageing imposes on lifespan.

Life as a continuity

We all recognize a living organism. This is thanks that we humans have the ability to recognise ourselves as such. This is something not quite obvious though, just a handful of species are recognized with the ability of mirror self-recognition (MSR) and the meaningfulness that this might
mean for self-awareness is still an open debate (de Waal, 2019). We humans do not only have the ability of MSR, but to do much more complex achievements such as creating and recording knowledge. In the particular field of life sciences, we have the ability to create biological knowledge. This biological knowledge should allow us to define, in a clear and precise manner, what is the fundamental definition of a life being. In short: we should be able to provide an indubitable answer to the question “What is life?”. A definition of life that would allow us to identify an alien form of life should we ever be able to find one. However, to date there is not yet a consensus answer to this question (Koshland, 2002; Benner, 2010; Tirard et al., 2010; Abel, 2011; Currais, 2017; Vitas & Dobovišek, 2019; Tetz & Tetz, 2020).

One thing that is widely accepted is that the origin of all living organisms currently on earth can be traced back to a unique event. An overwhelming number of observations points towards a shared origin of all living organisms on earth. To name some: 1) the fact that common aspects are shared at the molecular level, including fundamental biological polymers: nucleic acid as genetic material and L-amino acids as components of proteins; 2) the general unit of biochemistry in terms of basic metabolic and energy processes; 3) the existence of a nearly universal genetic code; 4) the congruence of morphological and molecular phylogenies; 5) the discovery of homologous protein families in the three domains of life (bacteria, archaea and eukaryotes); 6) the correspondence between phylogeny and the palaeontological record and 7) the existence of numerous predicted transitional fossils. Some attempts have been done to show the existence of a common ancestor with a formal quantitative testing (Theobald, 2010; Penny et al., 1982), although an agreement about if this can formally be tested has not been achieved (Velasco, 2018; de Oliveira Martins & Posada, 2016; Koskela & Annila, 2012; Koonin & Wolf, 2010; Theobald, 2010; Penny et al., 1982). The idea of a common ancestry dates back to Darwin itself (Darwin, 1871), but the idea of continuity in life is even older. Rudolf Virchow introduced the famous third dictum of cell theory: "Omnis cellula e cellula" (Virchow, 1859), which implicitly contains the idea of continuity ("All cells come from cells").

Ageing and the discontinuity of life

At the evolutionary level, it seems quite clear that there is continuity throughout the history of life on earth, but this is not at all obvious when analysing biological processes at the individual level. The phenomenon of ageing seems to impose a discontinuity on almost all living organisms.

As it happens with a definition of life, a common definition of what is ageing is not yet agreed (Fulop et al., 2019), and many controversies are still in place (Marsellach, 2017). Anyway, there is a definition of ageing commonly used in the ageing research community: ageing is defined as an increase in the mortality rate as an organism progresses throughout its life cycle and as its
chronological age increases (Rose, 1994; McDonald, 2013). However, this definition is only a macroscopic phenotypic description of ageing. It should be noted as well that, as it refers to mortality (which definition depends on the definition of life, definition that is not yet a closed issue), this definition of ageing, is not independent of the definition of life considered. In summary, this definition of ageing only describes what an external observer sees about the phenotypic characteristics of living organisms that do age but tells nothing about the molecular determinants that drives ageing. Nine hallmarks have been proposed (López-Otín et al., 2013), but no complete consensus is agreed yet (Zhang & Gladyshev, 2020; Fulop et al., 2019; Fuellen et al., 2019; Marsellach, 2017; Gladyshev, 2016).

Since there is no consensus on what ageing itself is, it is also difficult to agree on how aging arose. It was initially thought that aging evolved after the origin of eukaryotes, but observations of aging phenomena in prokaryotes questioned this initial belief (Ackermann et al., 2007). Some proposals have been made to tackle these observations: i.e., ageing is possible in single-celled organisms thanks to the dilution of the damage that takes place during cell division (Gladyshev, 2016; Gladyshev, 2013), and/or ageing evolved in single-celled organisms thanks to mechanisms that differently sequestrated damage to one of the two dividing cells (Moger-Reischer & Lennon, 2019), creating in this way an incipient difference between the soma and the germline (as predicted by the disposable soma theory of ageing (Kirkwood, 1977)).

A coherent information-based conception of both phenomena: life and ageing

In nowadays ageing research, ageing is almost a synonymous or a derivate of the concept of damage and damage accumulation (Ogrodnik et al., 2019; Moger-Reischer & Lennon, 2019; Gladyshev, 2016). Something must be damaged for an organism to age and die (as a result of that damage).

In the last decade though, there has been a growing interest in the involvement of the epigenome in ageing. An astonishing observation of a high correlation of some epigenomic marks and ageing started all this interest (Gibbs, 2014; Horvath, 2013; Bocklandt et al., 2011). However, to date there is no clear agreement on whether the epigenome has a central role in ageing (is it the cause or the consequence?), and neither is a consensus on what the epigenetic clocks might mean, which are the role that methylation marks do play on ageing and health or even the involvement of them in the rejuvenation process (Gladyshev, 2020; Niehrs & Calkhoven, 2020; Bertucci & Parrott, 2020; Zhang & Gladyshev, 2020; Jiang & Guo, 2020; Field et al., 2018; Horvath & Raj, 2018). Epigenetics, although recognized as one of the factors implicated in ageing (López-Otín et al., 2013), is hardly considered the sole cause of aging in some studies (Marsellach, 2018; Ashapkin et al., 2017; Marsellach, 2017; López-León & Goya, 2017; Ocampo et al., 2016; Gibbs, 2014). In this
manuscript I develop in detail an information-based conception of ageing, previously proposed by this author (Marsellach, 2018), in addition to proposing a coherent and testable definition of what life is based solely on the concept of biological information. In my proposal, the notion of continuity is a fundamental concept for defining both life and ageing.

As a summary: there is life while the biological information, necessary to overcome lethal environmental barriers (linked to complex lifestyles), is still useful for such a purpose. Death is caused by the inability to overcome lethal environmental barriers due to 1) extrinsic lethality factors (i.e., accidentality that destroys the usefulness of the biological structures constructed with the biological information); or 2) intrinsic loss of the usability of the biological information to overcome the lethal environmental barriers. We call this later intrinsic factor of lethality (due to the loss of usability of the biological information) as the ageing process and is due to the way in which biological information is treated by the ageing living beings (see below for detailed development). According to the proposed model, ageing is overcome during intergenerational steps due to specific information retrieval processes (the rejuvenation process). Ageing, therefore, implies a discontinuity in the flow of biological information, while life continues in ageing species thanks to the continuous flow of useful biological information provided by the parents to their offspring during the rejuvenation process that takes place in the intergenerational steps.

To describe my proposal, I need to delve deeper into the history of genetics and epigenetics to identify some misconceptions that might have led us to the inability to find coherent definitions to describe life and ageing. This historical review has led me to propose a modified description of the central dogma of molecular biology, first defined by Francis Crick (Crick, 1970), as well as to propose a complete retrieval of the chromosomal theory of inheritance (Boveri, 1904; Sutton, 1903) to describe the intergenerational flow of biological information. The next sections of this article contain a development of all these concepts in detail.

2. Biological information and the battle for Mendel’s legacy

Since Mendel uncovered the laws of inheritance in hybrids (Mendel, 1866), the question of where and how the biological information is stored was an open question in the biological field. The rediscovery of Mendel’s Laws in the early 1900s independently by Hugo de Vries, Carl Correns and Erich von Tschemmack, and works made by prominent scientist William Bateson marks the start of genetic studies as the science to study heredity (Bateson et al., 1902).
The first clue about where the biological information could be stored came from the chromosomal theory in the early 1900s. Works by Walter Sutton and Theodor Boveri proposed to consider the chromosomes as the bearers of the Mendelian factors (Boveri, 1904; Sutton, 1903). This theory was finally fused with Mendelian laws by Thomas Hunt Morgan (Morgan, 1915), leading to the establishment of genetics as a new discipline, separated from other biological disciplines. This had as a side effect the development of classical genetics, with an abstract non-chemical concept of the gene, and therefore a decline in the interest in the molecular nature of the genetic material (Deichmann, 2004).

During the first half of the XX century, the main consensus in the scientific community about what could be the molecular nature of the genetic material responsible for the Mendel’s Laws was that the genetic material was made of proteins (Deichmann, 2004). The chromosomal theory of inheritance established the chromosomes as the bearers of the mendelian factors. At that time proteins and DNA were already identified as the main components of the chromosome. Proteins were considered a complex compound capable of encoding genetic information, while DNA was considered a simpler compound, and therefore was not the first choice (Deichmann, 2004). Two experiments, though, brought the attention to the DNA as the bearer of the genetic information: first, the paper by Oswald Avery, Colin McLeod and Maclyn McCarty, where purified DNA was shown to be able to transform a non-virulent strain of pneumococcus into a virulent strain (Avery et al., 1944), and second, the paper by Alfred D. Hersey and Martha C. Chase, in which they demonstrated that DNA was responsible for multiplication of bacteriophages (Hershey & CHASE, 1952). However, it was not only after the Watson and Crick proposal of a structure for the DNA, that the DNA was widely accepted as the source of the genetic information (WATSON & CRICK, 1953b; WATSON & CRICK, 1953a). Watson and Crick model gave a coherent explanation on how information was stored: a double helix, with twice the information, would allow the information to be stored, copied, and transmitted from one ancestor to two or more descendants.

Although DNA was clearly accepted as the source of genetic information, it soon became clear that the specific phenotype of a cell was not only due to its DNA content. In parallel to all the development of the genetics field, the epigenetics field started to develop (Allis & Jenuwein, 2016). Historically, the word “epigenetics” was used to describe events that could not be explained by the genetic principles. Waddington defined epigenetics as “the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being” (WADDINGTON, 1942; Goldberg et al., 2007). Even before the Watson and Crick model, some phenomena were clearly away from what it was expected just from the well stablished genetic principles. As examples: position-effect variegation (MULLER & Altenburg, 1930), transposable elements (McCLINTOCK, 1951), X-chromosome inactivation (LYON, 1961) and imprinting.
Together with this, the concept of genomic equivalence made clear that the DNA sequence of a given cell was not the only factor governing the phenotype that a given cell shows (GURDON, 1962; GURDON et al., 1958).

Proteins were later shown, as well, to be able to transmit biological information from one organism to another. In 1982, the identification of a protein as the “Infectious Particle” in the scrapie disease, showed this (Prusiner, 1982). Strikingly, the problems that both, characterization of DNA as the “Transforming Principle” and prion proteins as “Infectious Particle” had, were similar. In both cases they had to prove the absence of the “opposite” component: proteins or DNA respectively (Prusiner & McCarty, 2006).

Although, the previous example of prions as the “Infectious Particle” example, nowadays almost everybody identifies biological information with the DNA. DNA is indeed synonymous of biological information. To be more precise, DNA is considered the source of the genetic information of an organism, and all the biological information that this organism has is thought to be derived from its own genetic information (with the rare exception of the well-known maternal effects characterized in Drosophila (Dobzhansky, 1935)). Biological information then could have many faces: from the DNA sequences itself, the RNA sequences, the protein sequences, the modifications of histone marks in a given histone tail, or the ordered response to a stimulus (external or internal) that generates a biological pathway, just to name some, but, since the fierce discussion about the molecular nature of the genetic information was clearly won by the DNA (WATSON & CRICK, 1953b; WATSON & CRICK, 1953a), we all see this cellular component as the ultimate source for all the biological information contained in a given living organisms. In other words: all the biological information of a given individual is ultimately dictated by the genetic information contained in its own DNA. This later statement is the hidden assumption that underline all the “Genome Wide Association Studies” (GWAS). Those studies try to find a relationship from the phenotypic characteristics of an individual and its own DNA sequence.

The concept that all information contained in the DNA of an individual could give rise to all its phenotypic characteristics has its foundation in the central dogma of the molecular biology, as proposed by Crick 50 years ago (Crick, 1970) (from now on referred just as the central dogma). In recent years, the central dogma has been updated (Morange, 2009; Shapiro, 2009; Koonin, 2012; Koonin, 2015; Tan & Anderson, 2020), but basically no change has been made in the role of DNA as the ultimate carrier of all heritable information passed on to new generations. In recent years, however, there has been increasing interest in transgenerational epigenetic inheritance in the scientific research community, increasing the focus on heritable information that is not transmitted through the genome, but through the epigenome (Lange & Schneider, 2010; Prokopuk et al., 2015; Blake & Watson, 2016; Xavier et al., 2019; Liberman et al., 2019; Duempelmann et al., 2020;
On top of that, the GWAS studies, mentioned before, are part of a central problem nowadays in biology: the missing heritability problem (Manolio et al., 2009). This problem goes into the heart of the concept of the transgenerational transmission of the biological information. Several solutions have been proposed for this problem, among them that the epigenome could be one possible answer for the missing heritability problem (Bourrat et al., 2017). In this paper I defend a complete recovery of the chromosomal theory of inheritance as a main framework for the understanding of how the biological information flows from one generation to the following ones (all the information, the information contained in the DNA sequence plus the information contained in other components of the chromosome). This means that both, the DNA and the proteins bound to that DNA (or modifications of that DNA; altogether known as the epigenome) do transfer biological information from the parental generations to the following ones. In fact, since the beginning of the questioning of where the biological information was physically stored, there has been a dispute between nucleic acids and proteins as the real agents of biological information. Therefore, it was considered that either DNA or proteins were the carriers of biological information. This fact assumes that if one of the two contenders won the battle the other has zero contribution to the heredity. But it could simply be that this dispute was, in fact, a human artificially created dispute (Marsellach, 2018).

In summary, the central dogma should be updated to include all new relevant findings, but specially to account for non-genetic generation-to-generation information transfer. To discuss all this I develop in the following section an updated version of the central dogma with the focus on the epigenetic information and discuss the convenience to recover in full the chromosomal theory of inheritance to account for the intergenerational information flow.

### 3. The epigenome, the hereditary code that depends on the genome but that is inherited *per se* separately from the genome.

The epigenetic code

As explained above, the word “code” in biology leads mainly to think in DNA and the central dogma, and therefore to reduce the meaning of the word code to a sequence of either nucleotides (in the case of DNA and RNA) or sequence of amino acids (in the case of proteins). This is starting to be seen as a clear simplification, and many more codes are being recognized inside the biological world. Although ultimately all the information is encoded in the DNA sequence, the products of that
code are needed to produce the decoding tools used to read the code (Tan & Anderson, 2020). Therefore, the DNA alone is useless and needs from other components in order to be executed. Francis Crick’s statement that there is no route to transfer back information from proteins to DNA is useful just when one thinks in the word “code” as “a sequence of”, but it has to be recognized that proteins do affect DNA function through other codes.

Epigenetics, like genetics started as an abstract concept with no idea of the mechanism that underlies them (Allis & Jenuwein, 2016; Deichmann, 2004). As molecular biology developed, and epigenetic mechanisms were started to be known, it soon become clear that it represents a code for itself (Jenuwein & Allis, 2001). The epigenetic code though is not as simple code as the genetic one (which can be reduced to “a sequence of”) but a much more complicated one and that allow to store much more information on the same length of the DNA fibre (see Table 1 and Figure 1).

As an analogy, the use of an epigenetic code “on top” of a genetic code can be seen as a similar transition of the one that happened in computer programming with the introduction of object-oriented programming languages (OOP) compared to previous procedural programming languages (see Figure 2). As discussed earlier and nicely explained in by Change Tan and Eric Anderson (Tan & Anderson, 2020), even the simplest task to be done over the DNA needs from extra components to be executed. In this regard always there is a feedback between the DNA and their products.
Figure 1: The Epigenetic Layer of Information. The epigenetic information adds an extra layer of information (blue cylinder) to the sequence information contained in the DNA (red cylinder). Arrows represent the exported functional information that comes from the DNA (red) or the epigenome (blue). Note that adding an additional layer of information "on top" of the DNA information allows more functional information to be stored for the same length of DNA fibre.

The existence of a dual code, the genome and the epigenome, allows to achieve many more complex achievements. This has some parallelism with the OOP paradigm with allow many more complex achievements thanks to definitory aspects of the OOP paradigm like for example the code reuse. The analogy is not perfect as one can apply the class-object paradigm not only to DNA binding proteins (as plotted in Figure 2) but to the whole transcriptome and proteome. What makes the DNA binding factors (or to any biological component that affects DNA or biological components “on top” of the DNA) different from other players from the whole transcriptome or the whole proteome is that, by lying “on top” of the DNA fiber they create a new layer of recorded information that can be transferred to the progeny. On top of that, the recorded epigenetic information is not only restricted to “a sequence of” but could be of many more types such as which kind of message contains (the concrete epigenetic modification), where it lies in the genome (on top of which concrete DNA sequence), or even how this affects the location of the DNA fiber inside the 3D structure of the nucleus, to name some (see Table 1).

In summary, the existence of a dual coding layer: the genome plus the epigenome creates richer information repository around the DNA: its coding sequence (the genome) and its associated factors (the epigenome); see Figure 1 and Figure 2.
Figure 2: The OOP Epigenetic code. Comparison between the computer programming languages paradigms of procedural programming (left) and OOP programming (right). A schematization of the of both types of computer languages paradigms are plotted on the top part of the figure (top left: procedural programming schema; top right: OOP programming schema). A biological example is provided in the bottom part of the figure. Colour coding is used to reflect comparable elements. Note that, as describe in the main text, this proposal does not pretend to be an exact analogy. On the procedural part of the figure: functions ≈ genes are plotted in orange, while “global data” ≈ “genome products” are plotted in blue. In the OOP part of the figure: class ≈ gene are plotted in orange, while objects ≈ “epigenetic factors” are plotted in blue.

An updated version of the central dogma

In the Figure 3 I have plotted an updated version of the central dogma to account for the additional codes that exist inside biological information and that are not restricted to the genetic code as in the Crick’s version of the central dogma (Crick, 1970). This proposal just wants to emphasize the existence of additional codes (especially the epigenetic ones) and does not pretend to be a final word in relation to the flow of biological information inside and between cells. As mentioned before some other proposals have already been made to update the central dogma.
Figure 3: Central Dogma of molecular biology. (A) Classical schematisation of the Central Dogma of molecular biology proposed by Francis Crick (Crick, 1970). (B) A new schematisation of the Central Dogma of molecular biology which include the effect of epigenetic factors in the flow of information. Black lines stand for Watson-Crick based information. Solid lines stand for main pathways of information flow. Dotted lines stand for less frequent alternatives. DNA replication black line points only to active DNA for clarity purposes. Purple lines stand for epigenetically-coded information. For purpose of clarity a single colour was used, but as detailed in Table 1, many different epigenetic codes do exist. Epigenetic factors are arbitrarily divided in three colour coded categories: activators (green), other chromatin functions (blue) and repressors (red). This affects both proteins and non-coding RNAs, but it should not be limited to those components (there is no certainty that other kind of epigenetic components might exist). Adapted from Marsellach 2018 (Marsellach, 2018).

Chromosomal theory of inheritance

What the chromosomal theory of inheritance says is that the information that is passed from the parents to their offspring is the one contained in the chromosome, the whole chromosome. At the time it was formulated the term genetics did not even exist (Deichmann, 2004). Genetics started as a science in which an external observer, who has direct access to some phenotypical characteristics, tried to use them to deduce the underlying phenomena that can explain how they flow from one generation to the following ones. Later on, with the birth of the molecular biology, the interest was not only in the heredity of those phenotypic characters, but on how they are mechanistically achieved. Genetically based characteristics (DNA based) were easily identified thanks to the work of Mendel and followers (Morgan, 1915; Bateson et al., 1902; Mendel, 1866). The beauty of Mendel’s Laws and its astonishing correspondence with some early experimental observations made genetics a flourishing field with many successes. The discovery of the double strand structure of the DNA seemed to close the circle about how the biological information was passed from the parent to their offspring (WATSON & CRICK, 1953b; WATSON & CRICK, 1953a). Together with many more achievements in the molecular biology field this led to the concept of the central dogma and to almost uniquely identify the biological information with the information contained in the double
helix of the DNA. This means genetic information almost became synonymous of biological information.

An external observer, however, has only access to the phenotypic information, with no *a priori* knowledge of the mechanisms by which this is achieved. A phenotypic characteristic that is only genetically-coded, will nicely follow the Laws of Mendel. This is easy to detect in the case of genetically-coded phenotypic characteristic coded in just one *locus*, but much more complicated for more complex genetic scenarios. To date, however, there is no accepted proposal of which pattern of inheritance one should expect for an epigenetically-coded phenotypic characteristic.

I propose to recover the chromosomal theory of inheritance in full. This means that the biological information that the offspring get from their parents is the full information contained in the chromosome, the one contained in the DNA fibre (the genetic information) and the one contained in the factors and modifications that are linked to that genetic information (the epigenetic information). In the next section I discuss about how the epigenetic information is handled during the meiotic and developmental processes that lead to a new generation.

**Flow of epigenetic information between a parental cell and its meiotically derived descendants.**

The study of the transmission of epigenetic information through the meiotic divisions is a hot topic in current research. Since the interest for the transgenerational epigenetic inheritance has flourished there has been two kind of opposite observations: 1) from one side there seems to be a complete erasure of the previous epigenetic information during the gametogenesis and the development of the newborn organism. This is a phenomenon well studied in mammals, and include two main types of epigenetic reprogramming, the gametic epigenetic reprogramming (Hackett & Surani, 2013; Cowley & Oakey, 2012; Hill et al., 2018), and the embryonic epigenetic reprogramming (Smith et al., 2012; Kobayashi et al., 2012; Hirasawa et al., 2008; Kono et al., 2004); and 2) from the other side there are evidences of transgenerational epigenetic inheritance in a wide variety of species ranging from yeast, plants, worms or mammals (Perez & Lehner, 2018; Yu et al., 2018; Klosin et al., 2017; Miska & Ferguson-Smith, 2016; Heard & Martienssen, 2014; Becker & Weigel, 2012; Greer et al., 2011; Grewal & Klar, 1996). In recent years however insights from many studies suggest that there is not a complete erasure of previous generations epigenome, and so that the contribution of epigenetic information to the following generations seems to be something quite usual (reviewed in depth by Xavier and colleagues (Xavier et al., 2019)).

As for genetic information there seems to be a universal way to describe how the genetic information flow from one generation to the next (Mendel, 1866), this is not yet the case for the
epigenetic for the epigenetic information. With the aim of finding a universal way to describe how the epigenetic information flow from one generation to the following ones, this author found some misconception in the studies dealing with transgenerational epigenetic inheritance. First, unicellular eukaryotic organisms like yeast can have descendants by both an asexual cycle (mitotic) or a sexual cycle (meiotic), but this is not the case in multicellular eukaryotic organism, which have cells than give rise to its own cells (the soma), and cells that will give rise to the following generations after a process of gametogenesis that involves a meiotic division (the germline). In multicellular organism only the germline can transfer epigenetic information from then parents to the offspring. Examples of recent literature include cases in which meiotic epigenetic inheritance is referred as a known issue in yeast (Skvortsova et al., 2018) while citing yeast work dealing with mitotic epigenetic inheritance (Rusche et al., 2003). Meiotic epigenetic inheritance is certainly a known issue in yeast (Grewal & Klar, 1996; Yu et al., 2018; Ragunathan et al., 2015; Recht et al., 2006), but the level of meiotic epigenetic inheritance characterization is not as high as the characterization of mitotic epigenetic inheritance in yeast (O'Kane & Hyland, 2019). Some studies in higher organism use yeast data to compare or explain studies of transgenerational epigenetic inheritance without looking properly if those studies refer to mitotic or meiotic processes. This is like comparing the processes that happen in the somatic cells with the processes that happen in the germline of a multicellular organisms. A clear misconception. Secondly, since transgenerational epigenetic inheritance was uncovered a revival of Lamarckian-like evolutionary processes has gained attention, being still an open debate in the evolutionary and epigenetics field (Koonin, 2014; Deichmann, 2016; Koonin & Wolf, 2009; Loison, 2018; Danchin et al., 2019). A Lamarckian view of transgenerational epigenetic inheritance have associated that reprogramming processes happening in mammal's gametogenesis and development are a loss of information, given that no parentally accumulated information would be transferred from one generation to the following one if this is erased and rewritten. Overall makes that reprogramming is seen as a loss of information. But, for a proper epigenetic mark to accomplish its function it has to be present when this is needed (in timely manner). A given epigenetic marks (or enzymes that builds them), needs to be expressed at certain cellular types and/or in certain developmental processes. If a particular epigenetic mark is lost, after the host cell has lost the ability to re-create it, then an epimutation is generated (see Figure 4). Therefore, a reprogramming wave happening during gametogenesis or developmental processes is not a process of loss of information but a of gain of information process. A process that puts the right epigenetic information in the right place and at the right time.

In summary, reprogramming is process of gaining information, not a process of losing information. This is another misconception that some studies dealing with transgenerational epigenetic inheritance have. I have developed a proposal to explain, from a non-Lamarckian perspective, how
the epigenetic information might flow from one generation to the following ones (Marsellach, 2018). I describe the proposed model in the following section.

Figure 4: Reading and loss of epigenetic marks. (A) For a given epigenetic mark to be deposited in a certain place on the genome (where its presence is needed to produce a functional outcome) the needed products have to be expressed at the time of deposition. For simplification purposes only a gene is considered. (B) If a given epigenetic mark is lost, shown as an empty cylinder, in a cell type that has no ability to recreate the epigenetic mark again an epimutation is generated leading to a loss in the functional characteristic mediated by this particular epilocus.

4. The Laws of Mendel, Schrödinger's cat-like interpreted

Not all the approaches to transgenerational epigenetic inheritance deal with the problem with a Lamarckian biased approach. The epigenetic reprogramming is thought to be a process that corrects epigenetic errors accumulated in previous generations (Reik, 2007; Faulk & Dolinoy, 2011; Chen et al., 2006; Bruno et al., 2015; de Waal et al., 2012; McCarrey, 2014; Marsellach, 2017). This is the foundation of the model that I propose. During the epigenetic reprogramming there is a rewriting of information and a repair of accumulated epigenetic mistakes (or epimutations). In most cases, these errors are corrected, but some are overlooked and passed on to subsequent generations leading to transgenerational epigenetic inheritance. In other words, what is inherited through epigenetic information is mostly mistakes.

The lack of repair of some defects could be due to a limitation of resources (or time) to correct all errors, which would lead to a scenario where a certain error could be repaired or not just randomly. This would make that some given epimutations are repaired and some not (which ones are repaired, and which not is what happens randomly). In accordance with this, it is predicted that, in
cells with lots of epimutations a higher number of them would not be repaired; while, in the other way around, in cells with few epimutations it is more likely that all of them are repaired.

In an ideal scenario a cell with few epimutations entering meiosis would have most of them repaired. Therefore, for the phenotypic characteristic, coded by those particular epigenetic factors a non-Mendelian ratio will be observed (see Figure 5B left and middle panels). However, if a cell with lots of epimutations do enter meiosis, then, in some given epimutations repair would not happen. For the phenotypic characteristics coded by these given epiloci a Mendelian ratio will be observed (see Figure 5B right panel). In Figure 5B, two scenarios are contemplated for a meiotic epigenetic repair: full meiotic epigenetic repair (Figure 5B left panel), and half meiotic epigenetic repair (Figure 5B middle panel). This is done to contemplate all theoretically possible scenarios.

Figure 5. Schrödinger’s cat-like Laws of Mendel for the segregation of the epigenetic information. Schematic representation of a several meiotic processes. (A) the Laws of Mendel governing the segregation of the genetic alleles. (B) Three panels to outline the segregation of epigenetic alleles. In the examples showing the segregation of the epialleles, three scenarios are contemplated: 1) Full meiotic repair (left), 2) half meiotic repair (middle), and 3) no meiotic repair (right), of a given epiallele. Cells are represented as big circles, one pair of independent solid black lines inside those circles represents one pair of homologous chromosomes (just one pair for simplification purposes). Sister chromatids, once DNA has been replicated, are plotted as joined solid black lines. Pairs of homologous chromosomes are joined by a dotted black line. Genetic DNA alleles are represented as red (defective) or green (wild type) dots, inside the solid black lines (to show that they are part of the DNA fibre). Epigenetic factors are plotted as small green circles. The absence of a given epigenetic factor in one locus are plotted as red dotted circle lines instead of the epigenetic factor. Both epialleles are plotted on top of the solid black line (the chromosome), to show that they do not strictly form part of the DNA fibre. Adapted from Marsellach 2018 (Marsellach, 2018).

As seen in Figure 5 this leads to a Schrödinger’s cat-like scenario when an external observer analyses the pattern followed by an epigenetically-coded phenotypic characteristic (see Figure 5). In comparison with the nice and easy Mendelian ratios followed by genetically-coded phenotypic
characteristics (see Figure 5A), an epigenetically-coded phenotypic characteristic would follow Mendel’s law only when the epigenetic mistake is not properly corrected (see Figure 5B).

In summary, the randomness nature of the epigenetic repair program happening in all meiotic cells might lead to a Schrödinger’s cat-like scenario when studying the segregation of the phenotypic characteristics coded by epigenetic factors.

5. Life and ageing as information-based phenomena.

The model

As detailed earlier, an increasing interest in ageing and epigenetics has recently appeared in the scientific research community. Ageing has been traditionally linked to damage, as if it was a wear and tear phenomenon. Recently, some works have proposed epigenetics as a unique cause of ageing (Gibbs, 2014; Ocampo et al., 2016; López-León & Goya, 2017; Marsellach, 2017; Ashapkin et al., 2017; Marsellach, 2018), but a formal probe of that is still missing. Below I propose a testable model that could help solve the epigenetic nature of the ageing process, and, at the same time, give a coherent description of what life is.

I propose to conceive ageing solely as an information-based phenomenon. In concrete as circular reuse of epigenetic information (Figure 6A). This proposal has two premises: first to conceive the ageing process as a loss of epigenetic information process. The epigenetic information inherited from the meiotically derived parental cells have the right information to produce a young phenotype. During the lifespan of a multicellular organism both the soma and the germline accumulate epimutations. The somatic cells will never read again the relevant epigenetic information for achieving a young phenotype (the “young epigenome”), while the germline cells will do so when producing the following generation. In the germline the accumulated epimutations are repaired and the relevant epigenetic information is put in the right place to achieve a “young epigenome”. A second premise of the proposed model is that the difference between these two scenarios is to continue alive (in the germline by receiving the appropriate functional information) or to die (in the soma due to a lack of adequate functional information due to loss of epigenetic information and the accumulation of epimutations). This later premise implies that lethality due to intrinsic factors is mainly epigenetically caused.
From the previous paragraph one can infer the proposal that I am doing to describe what is life: if you read the right information you will survive (the descendants of the germline) otherwise you will die (the soma). From a human perspective this is a nonsense: the soma accumulates epimutations that know how to repair, germline does it, but somatic cells do not do it because the tools for doing so are simply not used in those cells. This has a severe consequence: death of the somatic cells (and germ-line cells not emancipated from the organism). Why could this happen? Why dying if you have the right information for not to do it! The answer to this question is that there is a trade-off between complexity lifestyle and lifespan. The less complex your lifestyle, the less likely you are to die from an extrinsic factor and therefore you are more able to “afford” a long lifespan. The opposite is as well true, the more complex your lifestyle, the more chances to die from an extrinsic factor and therefore having biological mechanism that provide you a long lifespan would not increase your chances to live longer.

As detailed in the sections above, the acquisition of an epigenetic code during evolution might have allowed to achieve much more complex biological structures (there is more information available and more ways to bring this information into functional outcomes). Therefore, the more complexity was accumulated during evolution through the epigenetic code, the more complex lifestyles were developed. This had as a side effect the increase of chances of dying due to extrinsic factors (and therefore not to keep transferring the accumulated information to the offspring). The complexity-lifespan trade-off due to extrinsic mortality causes might had reached a point in which a further complexity acquisition did not increase the chances to keep storing new information (complexity-
lifespan trade-off limit; see Figure 7A). All organism above a certain level of complexity would probably die due to extrinsic factors and therefore not contribute further to the pool of living organism and neither to the pool of accumulated information.

Understanding ageing as an epigenetic unique phenomenon does not mean that ageing was an event that suddenly appeared in the evolution, but probably a slow evolving process, the transition from a pre-full-ageing world to a full ageing world was not an all or nothing transition from one day to the other, but a slow increase in the acquisition of non-genetic information and the use of it. However, with the appearance of eukaryotic organism and sexual reproduction (Speijer et al., 2015) the ageing process was mastered leading to a full ageing era (Marsellach, 2017). The appearance of sexual reproduction created an incipient distinction between the soma and the germline. In unicellular organism this was achieved inside the same cell by separating two kinds of divisions, the mitotic divisions (which did not lead to a sexual cycle) and the meiotic divisions (which led to a sexual cycle). This incipient distinction between the soma and the germline created two different paths: the path of epigenetic information recovery (in the germline-like cells) and the path of irremediable epigenetic information loss (in the soma-like cells), which eventually leads to death. In order to analyze the pathways that allow viability keeping in organism (either eukaryotic or prokaryotic) one has to differentiate between unicellular and multicellular organism.

**Unicellular organism**

Do unicellular prokaryotes age? What has this in common with unicellular eukaryotic ageing? An answer for this question could be, again, a complexity level related answer. Unicellular organism divides by mitosis (prokaryotes) or by mitosis and meiosis (eukaryotes), those organisms though, are far less complex than multicellular organism which needs from a developmental process to be shaped. The multicellular organism therefore needs much more information (either genetic or epigenetic) for its proper development and function than unicellular organism. A process of epigenetic information loss (as the ageing process is proposed to be in this essay) would then be affected by how many epigenetic information the organism has and how and for what is this information used for. In other words, ageing should be affected by the functions that the organism has to achieve via its epigenetic information.
Figure 7. Complexity-lifespan trade-off and lethality avoidance strategies. (A) Complexity-lifespan trade-off schematization. Likelihood of complexity increase is represented with respect to time. During pre-full-ageing era (left axis) this likelihood reached the value of zero due to increased chances of extrinsic death as complexity increases. In full ageing era (right axis) the mastering of ageing in eukaryotic cells (cyclical reuse of epigenetic information) allowed a further increase in complexity acquisition despite further increases in extrinsic lethality. In both cases a linear increase is plotted for simplicity purposes. No value (other than >0) is plotted at the starting of the ageing era as formal quantification of those phenomena is out of the scope of this paper. Complexity-lifespan trade-off limit line is plotted in red dotted line. (B) and (C), Lethality avoidance strategies of unicellular (B) or multicellular (C) organism. (B) An example of grow of unicellular organism in laboratory test tubes is shown. Epi-wild-type individuals are plotted in orange, while epimutants are plotted in red. (C) A schematization of the overlapping generation mechanism of lethality avoidance used by multicellular organism. Individual lifespans are plotted in quadrants shaped gradients that evolve from blue (epi-wild-type) to red (epimutants) organism. Lethality of epimutants occurs when they reach the temporal black line. Blue arrow represents transgenerational information transfer that gives an epigenetic young phenotype (epi-wild-type) to newer generations. This figure was, in some parts, adapted from the “Test tube true C.png” file created by Theresa Knott and obtained via Wikimedia Commons.

Unicellular organisms do not have complex developmental processes in their life cycle and use mainly epigenetic processes to deal with environmental changes (i.e., to switch from a nutrient rich media to a nutrient poor or and stressful media (López-Maury et al., 2008)). Epigenetics does not describe a developmental program as it does in multicellular organism. In this scenario, epigenetic information loss, and therefore the ageing process, might be reduced to a random accumulation of epimutations. Unicellular eukaryotic cells would survive through cycles of only mitotic divisions thanks to the ability to produce millions of identical copies of themselves. Among all those copies, the ones that have less epimutations accumulated (by chance) will overgrow the more epimutated ones (see Figure 7B). It should be noted that, the conditions in which for example yeast is grown in the laboratory, is far from the scenario that they find in the nature. It is probable then, that although they are able to grow for many generations under mitotic cycles in the rich media used in labs, in the nature they require a frequent passage through a sexual cycle for a prolongate survival. Unicellular procaryotes like bacteria and archaea, probably are still trapped in a pre-full-ageing era where viability is maintained by a more or least constant reading of the relevant information for survival (no distinction between soma and germline) and for the ability to quickly produce millions of descendants from just one singe organism (see Figure 7B). In the microbial world one can see ageing like scenarios given that all organism uses epigenetic factors (as detailed above DNA alone is useless without tools that decode it (Tan & Anderson, 2020)). As proposed by unicellular eukaryotes these pre-full-ageing phenomena should probably by random processes that asexually dividing organism do avoid by producing millions of descendants, by re-reading constantly the epigenetic information lost, or by a combination of both processes.

Multicellular organism

In multicellular organism, the full distinction between the soma and the germline allowed the appearance of a specific program to rewrite the epigenetic information in the germline so the functional information is recovered and passed to a newborn individual, which not only starts its existence with the youngest epigenetic information, but as well with its chances of dying by extrinsic factors reset to zero. Before its “conception” this organism didn’t even exist. Neither did its
chances of dying by an extrinsic factor. This is how, by overlapping generations of mortal individuals, the full development of the ageing process allowed to overcome the limitation of complexity acquisition due to extrinsic lethality factors (see Figure 7A and 7C). By overcoming the lethality linked to complexity limit, mortal organism, can keep accumulating and storing genetic and epigenetic information that allow a further increase in complexity.

Multicellular organism has at least one program to follow the development from the zygote to the newborn organism. On top of that epigenetic information might regulate further developmental stages after birth (see Figure 6B). The intense use of epigenetic information for developmental processes could have led to a program-like process that affects the loss of epigenetic information process in multicellular organism. In other words, in multicellular organism ageing is probably more a program-like process than a random-like process of epimutations accumulation (as proposed in this essay for unicellular organism), although it might be a mixture of both, a program-like process and a random (environmentally affected) process.

Altogether in all scenarios (unicellular or multicellular organism) according to the model presented in this paper, ageing is a consequence on how information is handled during vegetative grow and during the meiosis (and subsequent steps). During vegetative growth the epigenetic information is lost, while during processed happening in germline derived cells the relevant epigenetic information is read and used to create a young newborn organism. Ageing is merely a consequence of how epigenetically coded information is inherited.

The present and potential experimental evidences for the proposed model

The above proposed model is so far a speculation based in two premises (see above). In this section I summarize the experimental evidence already obtained to support this model, as well as I list some easy-to-check experimental frameworks that could help clarify whether the proposed model is accurate or not.

Experimental observations in accordance with the proposed model

The ageing field is starting to be hit by an epigenetic reprograming revolution. Through epigenetic reprograming some never before thought achievements are being accomplished, mainly by ectopically inducing reprograming via the four Yamanaka factors: Oct4, Sox2, Klf4 and c-Myc (OSKM) (Takahashi & Yamanaka, 2006). From resetting ageing hallmarks of cells from centenarian
individuals (Lapasset et al., 2011) to ameliorating age associated hallmarks in aged organism in vivo (Ocampo et al., 2016; Sarkar et al., 2020; Rodríguez-Matellán et al., 2020; Lu et al., 2020). All these new scenarios are difficult to explain with most of the long-time proposed models of ageing, especially the damage-based models of ageing, but are very easily explainable by the model presented in this paper: ageing as an epigenetic information-based phenomenon and epigenetic reprograming as a regain access to an always kept information. This information is never read again in somatic cells but read again during meiotic and developmental processes that start from germline cells to give rise to the newer generations. Ectopic reading of the relevant information can be induced ectopically in somatic cells by cell reprograming strategies (Takahashi & Yamanaka, 2006), leading to a rejuvenation of markers of ageing in those cells (Horvath, 2013). In physiological conditions though this reading of relevant information only happens in germline derived cells.

These great achievements have been preceded by almost a decade of increasing involvement of epigenetic processes in ageing (Zhang et al., 2020). This includes, among many more studies, the thriving field of epigenetic clock development (Liu et al., 2020; Ryan, 2020; Bell et al., 2019; Horvath, 2013; Bocklandt et al., 2011), and its relationship to all-cause mortality in humans (McCrory et al., 2020; Marioni et al., 2015). The relation between all-cause mortality and epigenetic clock signature suggests a somehow epigenetically caused mortality, as predicted by the model described in this paper. According to the model, lethality is both epigenetically caused and the end of the further flow of information (which, as proposed, describes life).

Altogether all these studies relating epigenetics and ageing favours an information-based conception of ageing rather than a damage-based conception of ageing. It has already been shown that modifying the epigenome (through reprograming) does reduce the impact of damage accumulated to the cells (Sarkar et al., 2020; Lu et al., 2020). This directly suggest that damage is not the cause of ageing but one of its consequences.

Another observation that is difficult to accommodate in a damage-based conception of ageing is the fact that Hutchinson-Gilford progeria syndrome (HGPS) is caused by a single point mutation in the Lamin A gene (Wang et al., 2020; Eriksson et al., 2003). This observation is difficult to explain in a damage-based conception of ageing, where damage accumulation occurs during lifespan due to changes into balance between damage generation and damage repair, and the consequent accumulation of damage that this implies all along the lifespan of the organism. Could all these complex mechanisms be caused by a point mutation? An information-based conception of ageing offers a more parsimonious explanation for such observation. As proposed, eukaryotic cells mastered the ageing process that had already begun in prokaryotic cells due to the incipient differentiation of the soma and the germline with the implementation of a meiotic division. The founding defining characteristic of eukaryotic cell though is not the meiosis, but the nucleus
surrounded by a nuclear membrane. The Lamin A protein is an integral part of the nuclear membrane, and by lying there could contribute to the spatial organization of the chromatin, that as proposed, could be an epigenetic information container (see Table 1). A single point mutation in Lamin A protein could therefore imply a sudden disorganization of nuclear architecture, characteristics that has been shown to happen in ageing and in particular to HGPS affected cells (Cenni et al., 2020; Reddy et al., 2020; Liu et al., 2011), and creating a quick information loss that could explain the premature manifestation of ageing phenotypes in HGPS patients.

At last, it should be noted that, as implicit with the model, the rejuvenation information should be kept silenced in the somatic cells. Indeed, meiotic genes has been shown to be specifically silenced in somatic (or somatic-like) cells (Hiriart et al., 2012; Harigaya et al., 2006). In germline (or germline-like) cells the old epigenetic information should be substituted by the young epigenetic information, so that, as shown in ectopically induced scenarios (Sarkar et al., 2020; Lu et al., 2020), decrease (or erase) the aged phenotype. This has already been shown to be the case in physiological conditions (Unal et al., 2011).

**Easy-to-check experimental frameworks for the proposed model**

The model proposed in this paper offers a possible explanation for the differences observed in lifespan duration in several examples based on the trade-off between lifespan and complexity (or extrinsic mortality likelihood): 1) most longest-lived organism known are organism showing a low-risk lifestyle (i.e., comb jellies, sponges, placozoans, hydrams, jellyfish, corals, sea anemones, myxozoans, planarias, acoels, red sea urchin, etc.). In those animals the low extrinsic lethality linked to their lifestyles could allow that they have a high non-meiotic reading of “young information” (or rejuvenation) leading them to extends its lifespan by delaying the epigenetic-mistakes-based lethality linked to ageing; 2) In organisms with highly efficient protective elements like turtles or clamps, or animals with low risk of predation due their constitutions (i.e., elephants, whales, sharks, etc.). In those cases, a slower speed of the ageing program or a higher non-meiotic rejuvenation could as well give rise to longer lifespans due to low extrinsic mortality likelihood; and 3) A lifestyle difference could as well explain big lifespan differences between phylogenetically related animals (i.e., long lived rodents like naked mole rat compared with shorter-lived rats). Experiments could be designed to test (i.e., via epigenetic clocks) how epigenetic drifts are affected in all those animals in comparison with short-lived examples.

At last, this author detected clues of epigenetically caused lethality in *S. pombe* (Marsellach, 2017). The framework that I developed allowed to put on test many predictions derived from the proposed model, in especial could allow to test whether life and ageing are indeed information-based phenomena by for example uncoupling meiosis from next generation production (Marsellach,
Those experiments, with a careful monitoring of information lost/gain could allow to put on test the above proposed model to describe life and ageing.

6. Conclusion

This author came up with this theoretical framework from unforeseen and incomplete observations (Marsellach, 2017). The model shows up from these partial observations by asking if they are the exception or the rule. Answering that those partial observations are the rule rather than the exception allows to build the proposed theoretical framework (Marsellach, 2018). Anyway, the model is coherent enough independently of the partial observations obtained by this author, and can be built by itself just by the paradigmatic shift of considering ageing as an information-based phenomenon rather than a damage-based one. In other words, by considering damage as the consequence of ageing rather than the cause of it. On top of that, the model is not only fully testable (as detailed above) but has already experimental support by data obtained by Sarkar and colleagues and by Lu and colleagues which shows that by reprogramming aged tissues they could obtain a measurable reduction of the already accumulated damage (Sarkar et al., 2020; Lu et al., 2020). These studies strongly support than damage is a consequence of ageing rather than its cause and provide evidence for an information-based nature of ageing.

In summary, the model proposed in this paper provides both a definition of life and ageing that are complementary and coherent with each other, and, more importantly, are testable. To the best of my knowledge this is a first proposal for an experimental test aimed at answering what life is. Based on all the considerations included in this paper, I would like to encourage the scientific community to consider the information-based nature of ageing as a plausible and more parsimonious explanation for the ageing phenomenon in their future studies, and to consider life and biological sciences as mainly information-based sciences. Biology is all about information handling.

Avery, O. T., Macleod, C. M., & McCarty, M. (1944). STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. 79(2), 137-158. /pmc/articles/PMC2135445/?report=abstract


https://doi.org/10.1186/s13062-015-0084-3


https://doi.org/10.1186/1745-6150-7-27

https://doi.org/10.1186/1745-6150-5-64

https://doi.org/10.1126/science.1068489

https://doi.org/10.3390/genes3010081

https://doi.org/10.1002/bies.201000030

https://doi.org/10.1101/gad.173922.111

https://doi.org/10.1016/j.conb.2019.09.012

https://doi.org/10.1038/nature09879

https://doi.org/10.1111/acel.13229


https://doi.org/10.1159/000477209

https://doi.org/10.1038/nrg2398

https://doi.org/10.1016/j.cell.2013.05.039

https://doi.org/10.1038/s41586-020-2975-4

https://doi.org/10.1038/190372a0

https://doi.org/10.1038/nature08494


https://www.biodiversitylibrary.org/item/27295


