PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Opinion piece

Cite this article: Milani L, Ghiselli F. 2019 Faraway, so close. The comparative method and the potential of non-model animals in mitochondrial research. *Phil. Trans. R. Soc. B* **375**: 20190186.
http://dx.doi.org/10.1098/rstb.2019.0186

Accepted: 16 September 2019

One contribution of 15 to a theme issue 'Linking the mitochondrial genotype to phenotype: a complex endeavour'.

Subject Areas:
evolution, developmental biology, genomics

Keywords:
comparative biology, non-model organisms, mitochondria, generalization, idiographics, nomothetics

Authors for correspondence:
Liliana Milani
e-mail: liliana.milani@unibo.it
Fabrizio Ghiselli
e-mail: fabrizio.ghiselli@unibo.it

THE ROYAL SOCIETY PUBLISHING

Faraway, so close. The comparative method and the potential of non-model animals in mitochondrial research

Liliana Milani and Fabrizio Ghiselli

Department of Biological, Geological, and Environmental Sciences, University of Bologna, Bologna, Italy

LM, 0000-0001-5052-2075; FG, 0000-0002-1680-8616

Inference from model organisms has been the engine for many discoveries in life science, but indiscriminate generalization leads to oversimplifications and misconceptions. Model organisms and inductive reasoning are irreplaceable: there is no other way to tackle the complexity of living systems. At the same time, it is not advisable to infer general patterns from a restricted number of species, which are very far from being representative of the diversity of life. Not all models are equal. Some organisms are suitable to find similarities across species, other highly specialized organisms can be used to focus on differences. In this opinion piece, we discuss the dominance of the mechanistic/reductionist approach in life sciences and make a case for an enhanced application of the comparative approach to study processes in all their various forms across different organisms. We also enlist some rising animal models in mitochondrial research, to exemplify how non-model organisms can be chosen in a comparative framework. These taxa often do not possess implemented tools and dedicated methods/resources. However, because of specific features, they have the potential to address still unanswered biological questions. Finally, we discuss future perspectives and caveats of the comparative method in the age of 'big data'.

This article is part of the theme issue 'Linking the mitochondrial genotype to phenotype: a complex endeavour'.

1. Model organisms

> For a large number of problems there will be some animal of choice or a few such animals on which it can be most conveniently studied
>
> —August Krogh, 1929 [1, p. 247]

In August 1929, at the Thirteenth International Physiological Congress in Boston, the Danish scientist Schack August Steenberg Krogh delivered a lecture entitled 'The Progress of Physiology' [1]. Krogh touched on many significant points—still relevant to this day—but the focus of his talk was the need for physiology to undertake a comparative approach to study processes in all their various forms across different organisms. In this context, Krogh emphasized the importance of choosing the right model for a given biological problem. As an example, he cited his former teacher, Christian Bohr (father of the famous physicist Niels Bohr), that used a tortoise to study the respiratory mechanism of the lung because its anatomical features facilitated the experiments (see [2] for a review). A postdoctoral assistant attended Krogh's lecture, and was deeply inspired by it. His name was Hans Adolf Krebs, who discovered the citric acid cycle (also known as Krebs cycle or tricarboxylic acid cycle). Krebs was so inspired that 46 years later wrote a paper entitled 'The August Krogh principle: for many problems, there is an animal in which it can be most conveniently studied' [3]. At the time of Krogh's lecture, Krebs was working in the laboratory of Otto Heinrich Warburg who had great success by choosing the right models, such as the sea urchin to study the changes in energy metabolism upon fertilization, the unicellular alga *Chlorella vulgaris* to study photosynthesis, and rats to study the metabolism of cancer [3]. In particular, the latter studies led to the discovery of

the Warburg effect (cancer cells tend to favour glycolysis over oxidative phosphorylation (OXPHOS), even in aerobic conditions). Krebs himself applied the Krogh principle by choosing the pigeon breast muscle to study the pathways of oxidative metabolism. A causal correlation between great accomplishments in life science and well-chosen model organisms is undeniable, and has been acknowledged for a long time. However, given the complexity of biological systems and the astonishing diversity of life, it is important to use models carefully, to fully exploit the advantages of the reductionist approach without falling into over-extrapolation.

What is a model organism? According to the definition by Levy & Currie [4, p. 328], a model is 'a cognitive stand-in: instead of investigating the phenomenon directly, one studies an easier to handle alternative'. In biology, the term is used in the context of mathematical modelling and computational simulations (theoretical models), and referred to particular organisms (model organisms). First, in its most familiar form, a model is a modified theoretical analogue that retains some features of a phenomenon under study, while simplifying others; analysing a simplified surrogate based on a mathematical construct it is possible to draw conclusions about a real-world target system. In the second acceptation, a model organism is a representative specimen of another organism or of a broader biological group, and inferences from such models are empirical extrapolations [4]. Jessica Bolker [5] classified model organisms in two types: surrogate and exemplary models. This distinction is important because it entails quite different criteria regarding the choice of the model, and the different kinds of conclusions that can be drawn from studying it. On the one hand, surrogate models are mostly used in biomedical research to understand the mechanism of human diseases and to develop treatments, but the same approach is also applied to commercially relevant species and for the conservation of endangered species (e.g. assisted reproduction). A fundamental assumption that is made when using such type of models is that the surrogate (e.g. the mouse) will respond to manipulation in the same way as the target (e.g. humans), in other words, we assume that the biological process we are investigating in the surrogate model is shared with the target, and we translate the findings from surrogate to target by direct substitution. Accordingly, the surrogate does not need to be phylogenetically close to the target, the only important thing is that the surrogate matches the target in regard to the investigated biological process [5]. On the other hand, exemplary models are used in basic science as representatives of broader groups with the purpose of understanding fundamental biological mechanisms and evolutionary processes. In this case, the results obtained are used to draw general conclusions based on inductive reasoning—that is making broad generalizations from specific observations—and the phylogenetic context is important, because it allows us to evaluate the representativeness of the model in a wider context with respect to particular traits [5].

The two approaches have been extremely successful, but it is important to issue caveats and limitations associated with both of them. As evolutionary biologists—and given the evolutionary framework of this Issue—we are focusing on exemplary model organisms. In this context, the most important concept that we need to keep in mind is that inductive reasoning is a powerful method, but can lead to false conclusions even if all the premises are true [6]. We discuss the caveats associated with generalization in the following section.

2. The problem with overgeneralization

Generalization is a double-edged sword. Inference from model organisms has been the engine for many discoveries in life science, but indiscriminate generalization leads to oversimplifications and misconceptions. It is quite common—both in specialized publications and even more so in the mass-media—to find titles of articles that omits the model organism from which the results were obtained. In this way the pairing between the model organism and the taxon that the model is representing becomes invisible [5]. The scientific community is well-aware of such misbehaviour, as proven by the popularity of the Twitter handle @justsaysinmice run by James Heathers, a data scientist at Northeastern University in Boston, Massachusetts. All that the account does is to retweet a news story or press release containing a catchpenny headline or bold statements, and adding the words 'IN MICE' to it. The account was created in April 2019, and now (just three months later) it counts nearly 66 000 followers. What initially started as a joke is now generating debates about the misuse of generalization and inductive reasoning. To omit the subject of the study, especially (but not only) in biomedical research, generates the misconception that the finding is valid for every/most organisms, humans included. The bad practice of inflating research findings by overgeneralization can originate either at the source—owing to the increasing pressure for publishing 'striking and breakthrough' research in order to get funding or make a career—or during the process of promulgation to the public. The dissemination of wrong or distorted concepts is particularly bothersome because it undermines the trust of people towards science, or, even worse, creates anti-scientific mediatic monsters. There is an ongoing crisis of mistrust in science [7], and misinformation is playing a key role—albeit not being the only cause—in the decline of public confidence in scientists and experts in general. We think that the scientific community should take action to reverse the trend that has been pushing science in the direction of sensationalism and 'breaking news'.

During his life, Thomas Hunt Morgan—one of the fathers of genetics and responsible for making *Drosophila* a model organism—worked on a considerable number of different species [8], actually following Krogh's principle, nevertheless, he considered the results of research on fruit flies as applicable to all sexually reproducing organisms. Such perception regarding *Drosophila* and other model organisms is quite common among biologists and almost ubiquitous in modern textbooks [4]. Francis Crick called his theory about the information flow from DNA to proteins 'Central Dogma' [9]. A dogma is a principle laid down by an authority as incontrovertibly true, which is a concept foreign to science. According to Sydney Brenner—who established *Caenorhabditis elegans* as a model organism—Crick regretted using the word 'dogma', even if it was a joke, because Crick was not religious and did not believe in dogmas [10]. Also Jacques Monod, in 1954, stated that 'Anything found to be true of *E. coli* must also be true of elephants'. Actually, the phrase was coined by the microbiologist Albert Jan Kluyver in 1926 [11], in any case, this is another famous example of an overstatement, it is only partially true.

The problem with generalization was well known to both Krogh and Krebs. Krogh considered 'a general physiology which can describe the essential characteristics of matter in the living state' as an ideal goal that might need a long time to be achieved, and only by studying the same processes/

structures in 'myriads of organisms' [1, pp. 246–247; 12]. In 1980, Hans Krebs, together with John Krebs, wrote a paper warning about the 'uncritical application' of the Krogh principle [13]. As they put it: 'This article is a postscript to the earlier essay. It makes the point that an uncritical application of the principle may lead to fallacious generalizations, because extrapolating findings from one species to another is not invariably valid [13, p. 379].' They also pointed out that while a 'unity of life' certainly exists, there is also an 'infinite diversity of life', based on different combinations of a small number of basic units. Some biologists are devoted to finding common features, others are devoted to finding diversity, and often the choice between being, respectively, a 'lumper' or a 'splitter' depends on the personal attitude towards the same scientific issues. Needless to say, both approaches are equally important, and need to be carried out together [14]. Biology is extremely complex, but while simplification is fundamental to understand complex mechanisms, we cannot underestimate diversity [5,12,14–18]. Diversity is both the motor and the result of the evolutionary process. Evolution 'overcomes' problems using whatever is available at the moment—a process described by François Jacob as 'tinkering' [19]—thus finding multiple solutions for a given problem. An astronomically high number of genotypes can produce the same phenotype, and all these genotypes are organized in networks connected by point mutations [20]. At the same time, the chances of overcoming a problem depend on what is available, namely from variability: contingency is an intrinsic property of evolution, so is diversity. For these reasons, it would be wise to avoid the terms 'dogma' or 'law' in biology.

Model organisms and inductive reasoning are irreplaceable: there is no other way to tackle the complexity of living systems. At the same time, it is not advisable to infer general patterns from a very limited number of species, which are very far from being representative of the diversity of life. How to solve this conundrum? August Krogh in 1929 pointed the way by endorsing the comparative method: 'study all sorts of organisms'. Some are better than others, depending on the problem under study. We must choose an organism because it has properties suitable for investigating a particular phenomenon, not as a representative for a wide taxonomic group or life in general. We must have 'myriads' of models.

3. 'Rosetta Stones' or aberrant creatures? The model-organism-based approach versus the comparative method

Model organisms are often chosen because of practical advantages. They are usually easy to find and breed, convenient to maintain in the laboratory, have a short generation time and rapid development, and respond well to experimental techniques and manipulations [5,17]. We will incorporate such and related features under the term 'accessibility', which is what makes a model system more tractable and allows the researchers to focus on specific biological questions, disentangling them from a complex background. Other than accessibility, when choosing a model system it is fundamental to assess its 'representativeness'—that is, the ability to serve as a basis for inference [5]—with respect to the biological feature that is investigated. At the dawn of molecular biology, bacteria and bacteriophages were used as model organisms to study DNA replication, transcription and protein synthesis. Because such processes and the molecules involved are shared by all organisms on this planet, the representativeness of these models is very high, and the choice turned out to be quite successful. In that context, bacteria and bacteriophages were 'Rosetta Stones'—namely, something that acts as a key to some previously unattainable understanding [14]. Moving on to studying more complex biological processes, the researchers needed to choose other model organisms, like *C. elegans*, *Drosophila melanogaster*, *Mus musculus*, *Rattus norvegicus* and others (see [21] for a detailed analysis). Of course, over time, as more knowledge accumulates, the life science community is going to study increasingly more complex traits, and this is where biodiversity kicks in, undermining the representativeness of the 'traditional' model organisms. As pointed out by Krogh, the solution is to find multiple models, that is, to explore diversity to have a complete picture of the subject of study. From a conceptual point of view, on the opposite side of 'Rosetta Stones', there are 'aberrant creatures' [22] or 'extreme organisms', that is, highly specialized organisms that show very peculiar features. Frequently, the traits that make an organism an attractive model are the result of adaptations, distinctive features that can be experimentally convenient and increase its accessibility. One good example is the work of Krebs on OXPHOS that was carried out on pigeon breast muscles because of the robustness of the mitochondria in this kind of tissue. The adaptation to flight made breast muscle mitochondria in birds particularly resistant to stress, thus to experimental protocols [18]. However, adaptations reduce the representativeness of a model, so there is often a trade-off between accessibility and representativeness [5], which needs to be taken into consideration when extrapolating findings from a model species to more inclusive taxa. It is interesting to note that while 'Rosetta Stones' are suitable to find similarities across species—even if genuine 'Rosetta Stones' for living systems, in general, do not exist—extreme organisms can be used to focus on differences. Such kinds of negative models can be exploited to understand, for example, plasticity and limitations of a biological process (e.g. naked mole-rats for cancer research, and snakes for physiological adaptation to starvation, see [18] for a thorough discussion).

(a) How to infer general concepts and the role of model organisms in comparative biology

When making inferences in biology, we have to pay particular attention to the kind of questions we want to address and to the taxonomic range we want to investigate: are we looking for uniqueness or for more general mechanisms? A significant contribution to the understanding of this dualism is the concept of idiographics versus nomothetics. These terms were used by the Neo-Kantian philosopher Wilhelm Windelband to describe two distinct approaches to knowledge [23,24]. Several disciplines—historical sciences such as anthropology, palaeobiology, evolutionary biology and evo-devo—embody both principles. Idiographics is the description of unique and historically contingent features, while nomothetics is the discovery of regularities that can be described as general rules, and concerns the formulation of general concepts. This dualistic perspective helps us to understand the role of model organisms in biological disciplines [25]. When we study a specific model organism, we can only deal with the idiographic range. It is

not correct to generalize when studying a single taxon, because the comparison between individuals of the same species can provide insights into the biology and evolution of the species itself, but not necessarily into the biology and evolution of its relatives or other distant taxa. Generalizations are easy to propose in biology but seldom valid across all organisms [26,27]. How can we translate unique features to regularities? Namely, how can idiographics be related to nomothetic conceptual categories? We can compare and generalize only if we have a rich idiographic foundation that documents the evolutionary change of the traits we are taking into consideration in our study. This is the reason why there is a need of a judicious choice of model organisms to mine for best idiographics [25]. What is the best choice? If we minimize character overlap between model organisms by maximizing phylogenetic diversity, we consequently maximize the amount of unique idiographic details captured by the models. Consequently, we minimize the ability to draw general conclusions from them. As we pointed out above, given that model organisms are usually chosen for specific traits, any general conclusion drawn from these species is rarely generalizable [28].

Even though analysing a broad phylogenetic range of taxa is fundamental to assess diversity, it may not be the best choice to generate nomothetic insights [25]. We can access nomothetic themes if we maximize character overlaps between model organisms, choosing them for specific shared traits, as part of a same class. A class is a group of biological entities (e.g.: organisms or parts of them, biochemical pathways, gene regulatory networks, etc.) that share features which can be meaningfully compared. Such a concept expands, or 'completes', the concept of homology, and allows for comparisons beyond monophyletic taxa [27]. A class can be a group of organisms that share common traits, even if independently evolved, thus distinguished from the taxonomic term 'class'. From here ahead, when speaking about classes, we refer to the non-taxonomic entity. This ahistorical formulation of classes provides the necessary basis to propose nomothetic insights. Idiographics are most efficiently translated into nomothetic insights when model organisms are chosen based on sharing certain traits, that is, as members of the same class [25,27]. So, models should be chosen for possessing traits that provide an independent support for a particular concept rather than for their phylogenetic position *per se*; by doing this, the explanatory force—namely, the extension of predictive reliability—can be maximized. Of course, the availability of robust and extensive phylogenetic hypotheses is still fundamental to understand the direction of evolutionary change and the conditions under which certain traits can evolve.

To understand how we can approach generalization, we have to consider the specific jurisdiction in which the outcome of a specific experiment can be included. We have to consider the 'experimental model-organism-based approach' versus the comparative approach. There is no *a priori* reason to generalize the outcome of experiments based on a single model organism, as a single counterexample actually falsifies any generalization [29]. Furthermore, this kind of approach cannot give direct evolutionary inference, because the experimental changes that we can induce in the laboratory are just an analogy to evolutionary changes: it is not correct to compare a newly obtained phenotype in one species with naturally occurring phenotypes in that species or in others, even if similar, and deduce that a similar change has indeed happened in the course of evolution [29]. Similarly, it is not correct to compare the performance and fitness of manipulated organisms with 'wild-type' organisms, because the latter are functionally balanced and probably mutations disturb this functional balance [29]. For this reason, a reduced fitness may not be a support for the evolutionary stability of characters: the proximate causes of functional disadvantages of change cannot directly explain the ultimate causes of evolutionary change [29]. Only comparison allows generalization.

Most of the research performed with model species has been justified by their potential power for understanding human biology. To trace general themes and eventually understand life, we have to overcome such limited view and expand research to 'new' model organisms from a wider taxonomic range and in a wider set of environmental conditions. The term model organism was initially used to describe 'an organism that is inherently convenient to study a particular area of biology'. When saying 'model organism', we often mean 'an organism for which a wealth of tools and resources exist' [30]. To date, several organisms are used as models particularly suited to address some research areas—thus being model organisms in the original sense, that is convenient for the study of a biological process—but these often do not possess implemented tools, and dedicated methods and resources. Russell *et al*. [30] refer to these organisms as 'non-model model organisms' (NMMO).

(b) An example of the use of classes: the link between mitochondria and germline formation

A clear example of the use of classes in building reliable generalizations regards germline development [31] and its link with the mitochondrial inheritance process. Looking at the majority of model species studied so far, the segregation of germline from somatic lineages appears to happen early. The so-called preformation mode—that is, early development controlled by maternal factors stored in the egg—seemed to be the predominant mode of germline specification. Preformation was described for *C. elegans*, *Drosophila*, *Gallus gallus* and plenty of other models. However, as data from additional animal taxa are collected, it is becoming clear that the most common and ancestral mode is epigenesis, characterized by a later segregation of germ cells owing to inductive signals in the developing embryo [31]. The first idea was the result of the biased choice of models which have in common a faster development. Actually, developmental speed is deeply connected to the use of maternal material in early development. If transcripts newly produced by the embryo and their products are necessary for development to proceed, of course, the process is slowed down. If we chose models with slow development, the chance of choosing animals developing through inductive signals would probably increase. If we consider the class of taxa sharing the preformation mode of germline specification, we may be able to identify the general features that can lead from the ancestral epigenesis to an early way of germline specification and understand how the segregation of material delivered to the progeny has evolved. This kind of approach can be applied to any kind of meaningful trait.

The relevance of germline formation goes well beyond the field of developmental biology. Investigating the times and modes of mitochondrial segregation into the germline is of basic importance: the mechanisms leading to the segregation of mitochondria into germ cells can vary depending on the type of germline specification, early or late [32]. By sampling

animals with one or the other mode, in distantly related taxa, we have the possibility to address the general principles that lead to one or the other way of specification and at the same time to find, for the two modes, the linked mechanism of mitochondrial selection [32].

Given the increasing knowledge about the role of mitochondria in germline formation, these two topics have become tightly intermingled. The presence in many distantly related taxa of mitochondria associated with the germ plasm—material found in germ cells and containing germline determinants—suggested a general mechanism: mitochondria release material for germline specification/formation. Many studies on several animals documented the emission of mitochondrial material merging with germ plasm; this was observed in model organisms such as *D. melanogaster*, *Hemicentrotus pulcherrimus*, *Xenopus laevis* and *M. musculus* (see [33,34], and references therein), as well as in non-model organisms (e.g.: [35]). Moreover, a mitochondrial ribosome-dependent translation in germline formation is required at least in *Drosophila* embryos [36]. However, despite this evident link between mitochondria and germ plasm, the mechanism of such interplay is still unknown.

4. Rising non-model model animals in mitochondrial research

The growing understanding of the diversity of life clearly advises against the routine of formulating hypotheses and inferring general patterns based on a small number of species. Historically, general conclusions inferred from studies based on a skewed and limited taxon sampling led to a restricted and incomplete knowledge of mitochondrial biology. This uneven sampling is reported in table 1 which shows the number of complete mitochondrial genomes available in the NCBI Nucleotide database, subdivided across animal phyla. The poor (or non-existent) sampling in some groups is evident, as it is the skewed sampling in other groups. For example, human mitochondrial DNA (mtDNA) represent the 76.7% of all the chordate sequences deposited in the database. More recent evolutionary research (see for example: [37–40]) produced clear evidence that some of the 'textbook notions' about mitochondria are to be deemed as inaccurate, or plainly wrong. A major consequence of this is an increased awareness of the central role of mitochondria in eukaryotic life that goes beyond the concept of 'powerhouse' [41–45]. Accordingly, Geoffrey Hill recently coined the term 'mitonuclear ecology' to indicate the emerging interdisciplinary field that studies fundamental concepts of evolutionary ecology such as sexual reproduction, the origin of two sexes, sexual selection, adaptation and speciation in the light of mitonuclear interactions [41,46]. Exploratory research on non-model organisms broadened the variability of mitochondrial biological features well beyond the 'textbook notions' [47]. For this reason, it is necessary to exploit the new technologies to extend the number of species under investigation, including new models and understudied groups. Such 'uncommon' biological systems can help unveiling unknown elements of mitochondrial biology and evolution. In this section, we discuss the potential of non-model animal systems in mitochondrial research and highlight some promising candidates. Moreover, this Issue includes research papers focused on non-model animals such as bivalve molluscs [48], electric fishes [49] and mayflies [50].

There are notable examples of NMMOs systems in mitochondrial biology. We enlist some representatives in table 2, including fishes, mammals, birds, crustaceans and bivalve molluscs. This is not intended to be a comprehensive review of the whole range of NMMO animal species under study to date, but it is just an example of how new model organisms can be chosen in a comparative biology framework. The NMMOs highlighted here have special, often unique, traits to address questions of general importance. These organisms are studied not only because they are a kind of 'aberrant creatures' useful to gather novel idiographic records, but they can also be fundamental system to address central questions about biology that have remained unanswered to this day [30].

(a) Killifish (*Fundulus* spp., *Nothobranchius furzeri*): environmental adaptation, ageing

Members of the genus *Fundulus* have been investigated for adaptation to different temperatures [51–53,55] and different salinity [57]. Differences within and between species (different populations of *Fundulus heteroclitus*, and *Fundulus grandis*) showed the involvement of mitochondria in thermal adaptation. Indeed, temperature changes affect metabolism and mitochondrial performance, playing a role in whole-animal thermal tolerance and plasticity. These studies support a general role for OXPHOS and mitochondrial oxygen kinetics in differentiating aerobic performance and in influencing species responses to environmental change [53]. Comparisons of mitochondrial responses to thermal changes in different tissues investigated using high-resolution respirometry in subspecies adapted to different temperatures documented an extensive plasticity in mitochondrial performance following thermal acclimation in killifish, with the extent of responses differing between tissues [52]. Also, lipid remodelling in mitochondrial membranes appears to be a mechanism contributing to these changes [54]. *Fundulus heteroclitus* is also studied for cytonuclear disequilibrium that appears to act in reproductive isolation or selection against hybrids [56]. The species is abundant in estuaries and it shows adaptations to different water salinity. Comparative experiments were performed to track the transcriptomic and physiological responses to salinity variation, and to explore the regulatory mechanisms that may enable osmotic acclimation. Some genes showed population- and salinity-dependent patterns of expression during acclimation, including genes involved in mitochondrial function [57].

The killifish *Nothobranchius furzeri* is becoming a model for the study of senescence, being naturally short-lived [60]. Shorter- and longer-lived individuals of the species differ in gene expression, especially in early life. Mitochondrial complex I expression resulted to be negatively correlated with lifespan, with the phenotype reversed by complex I inhibitors (reversion of ageing-related regulation of gene expression and extended lifespan) [58]. An integrative genomic and genome-editing toolkit using the *de novo*-assembled genome and the CRISPR/Cas9 technology was developed for this species. This approach will be important to target candidate genes related to ageing [108], including mitochondrial dysfunction [59]. Ageing in *N. furzeri* appears associated with a decline in mtDNA copy number, downregulation of mtDNA-associated genes and an impairment of mitochondrial function [60]. This species appears to be an ideal model to assess the role of physiological and environmental parameters on ageing and lifespan determination.

Table 1. Availability of complete mitochondrial genomes on the NCBI Nucleotide Database. (The table shows the number of living species for each animal taxon, as reported by the 2019 Annual Checklist of the 'Catalogue of Life' (CoL) project (https://www.catalogueoflife.org/col/info/totals), and the number of complete mitochondrial genomes available on the NCBI Nucleotide database (data collected on 7 July 2019). In the 'notes' column, we indicated strong biases towards a restricted number of species. Between brackets the percentage of mitochondrial genomes of the cited species, with respect to the total number of available genomes.

taxon	current number of living species in CoL	no. of complete mt genomes	notes
Acanthocephala	1330	7	
Annelida	14 399	100	
Arthropoda	1 082 297	4427	*Anopheles* spp. 393; *Drosophila* spp. 206; *Magicicada* spp. 125; *Daphnia* spp. 90; *Apis mellifera* 67; *Bombyx morit* 61. Total = 942 (21.3%)
Brachiopoda	396	5	
Bryozoa	5434	10	
Cephalorhyncha	237	3	
Chaetognatha	132	11	
Chordata	69 913	63 801	*Homo sapiens* 48 026; *Pan* spp. 239; *Menidia menidia* 190; *Camelus* spp. 179; *Mus* spp. 162; *Rattus norvegicus* 118. Total = 48 914 (76.7%). *Homo sapiens* = 75.3%
Cnidaria	11 151	237	*Acropora* spp. 72 (30.4%)
Ctenophora	200	5	
Cycliophora	2	0	
Dicyemida	122	0	
Echinodermata	6828	96	
Entoprocta	171	2	
Gastrotricha	852	1	
Gnathostomulida	100	2	
Hemichordata	139	6	
Micrognathozoa	1	0	
Mollusca	65 442	791	*Architeuthis* spp. 38; *Mytilus* spp. 38; *Potamogyrus* spp. 23; *Octopus* spp. 21; *Unio* spp. 20; *Anodonta* spp. 12. Total = 152 (19.2%)
Myxozoa	245	4	
Nematoda	3455	258	*Anisakis* spp. 17; *Caenorhabditis* spp. 17; *Ascaris* spp. 12; *Angiostrongylus* spp. 11. Total = 57 (22%)
Nematomorpha	361	0	
Nemertea	1371	19	
Onychophora	167	5	
Orthonectida	25	0	
Phoronida	19	1	
Placozoa	2	6	*Polyplacotoma mediterranea* 1; *Trichoplax aderens* 1; 'Placozoan sp.' 4.
Platyhelminthes	18 616	436	*Schistosoma* spp. 149; *Echinococcus* spp. 113; *Taenia* spp. 21. Total = 283 (64.9%)
Porifera	9092	66	
Rotifera	2014	3	
Sipuncula	205	7	
Tardigrada	1018	4	
Xenacoelomorpha	456	12	

Table 2. Non-model model animals in mitochondrial studies.

taxon	special features	investigated traits	references
class Actinopterygii			
Fundulus heteroclitus (Atlantic killifish)	inhabits a steep latitudinal thermal gradient, with variation in mitochondrial properties	role of mitochondria and oxidative phosphorylation in variation in the pace of life mitochondrial dysfunction	Baris *et al.* [51] Chung *et al.* [52] Chung *et al.* [53] Chung *et al.* [54] Fangue *et al.* [55] McKenzie *et al.* [56] Whitehead *et al.* [57]
Nothobranchius furzeri (African turquoise killifish)	currently, the shortest-lived vertebrate bred in captivity it recapitulates typical age-dependent phenotypes and pathologies	ageing	Baumgart *et al.* [58] Harel *et al.* [59] Hartmann *et al.* [60]
class Mammalia			
Peromyscus maniculatus (North American deer mouse)	inhabits high altitude has evolved an enhanced aerobic capacity in hypoxia	mitochondrial basis for adaptive variation in aerobic performance	Cheviron *et al.* [61] Cheviron *et al.* [62] Lui *et al.* [63] Mahalingam *et al.* [64] Nikel *et al.* [65] Scott *et al.* [66] Scott *et al.* [67]
Heterocephalus glaber (naked mole-rat)	high longevity cancer resistant rarely suffering from age-related diseases	ageing age-dependent structural and functional changes of mitochondria	Bakeeva *et al.* [68] Buffenstein [69] Dammann [70] Holtze *et al.* [71] Khayal *et al.* [72] Kim *et al.* [73] Yu *et al.* [74]
Myotis myotis *Myotis brandtii* *Myotis lucifugus* (bats)	the only mammals capable of powered flight, with an extremely high metabolic rate, but associated with extended longevity	longevity/ageing heteroplasmy challenging the free radical theory of ageing	Ball *et al.* [75] Brunet-Rossinni [76] Brunet-Rossinni & Wilkinson [77] Dammann [70] Huang *et al.* [78] Jebb *et al.* [79] Munshi-South & Wilkinson [80] Seim *et al.* [81]
class Aves			
Eopsaltria australis (Eastern yellow robin)	low nuclear differentiation with two highly divergent, parapatric mitochondrial lineages with sharp climate-correlated differences in their distributions	climate-driven mitochondrial selection mitonuclear coevolution	Lamb *et al.* [82] Morales *et al.* [83] Morales *et al.* [84] Sunnucks *et al.* [85]
class Hexanauplia			
Tigriopus californicus	populations with extreme mitochondrial DNA divergence	thermal tolerance and its relationship to mitochondrial functions mitonuclear coevolution	Barreto & Burton [86] Barreto *et al.* [87] Burton *et al.* [88] Edmands [89]

(*Continued.*)

Table 2. (*Continued.*)

taxon	special features	investigated traits	references
			Ellison & Burton [90] Ellison & Burton [91] Foley *et al.* [92] Harada *et al.* [93] Willett [94]
class Bivalvia			
species with doubly uniparental inheritance (DUI) of mitochondria	natural and evolutionary stable mitochondrial heteroplasmy transmission of sperm mitochondria across generations the longest-lived animal to date has DUI (*Arctica islandica*: the specimen 'Ming' was 507 years old; [95]	mitochondrial biology and inheritance heteroplasmy challenging the free radical theory of ageing longevity/ageing mitonuclear interactions genomic conflicts	Bettinazzi *et al.* [96] Bettinazzi *et al.* [48] Capt *et al.* [97] Ghiselli *et al.* [98] Ghiselli *et al.* [99] Ghiselli *et al.* [100] Ghiselli *et al.* [101] Iannello *et al.* [102] Milani [32] Milani & Ghiselli [103] Milani *et al.* [104] Milani *et al.* [105] Skibinski *et al.* [106] Zouros [107]

(b) Deer mice (*Peromyscus* spp.): environmental adaptation

At high altitude, small mammals have to support thermogenesis and aerobic exercise in spite of the reduced oxygen availability. To address this topic, many studies have been performed involving highland and lowland *Peromyscus maniculatus* [63,65]. The integration of data from different approaches, such as measures of whole-organism thermogenic performance, measures of metabolic enzyme activities and genomic/transcriptional profiles, was used to examine the adaptive variation of this complex trait in deer mice that are native to different elevations. Different species of the same genus have been compared for such research (*P. maniculatus* and *Peromyscus leucopus*) [61]. Highland deer mice have an enhanced thermogenic capacity under hypoxia compared with lowland conspecifics and closely related lowland species. This feature is largely owing to an increased capacity to oxidize lipids associated with elevated activities of muscle metabolic enzymes that influence flux through fatty-acid oxidation and OXPHOS pathways in high-altitude deer mice, and by concomitant changes in the expression [61]. Other than having higher respiratory capacities in high-altitude mice than in low-altitude mice, the former also showed higher mitochondrial volume densities [64]. The association between transcriptomic profiles and muscle phenotypes was also tested: several genes involved in energy metabolism were more expressed in highlanders, and the regulators of mitochondrial biogenesis were positively correlated with muscle oxidative phenotype [66]. In species that are distributed across steep environmental gradients, adaptive variation in physiological performance may be also owing to transcriptional plasticity in underlying regulatory networks [62].

(c) Naked mole-rats (*Heterocephalus glaber*): ageing

Naked mole-rats (*Heterocephalus glaber*) are the longest-living (approx. 30 years) rodents known to date [69], thus being an important model to examine mechanisms modulating ageing. Naked mole-rats show negligible senescence, no age-related increase in mortality and high fecundity until death [73]. In addition, they are resistant to cancer—both spontaneous and experimentally induced—thus challenging the theories that link ageing, cancer and redox homeostasis. Indeed, although characterized by significant oxidative stress—young individuals surprisingly have high levels of accrued oxidative damage [69]—the naked mole-rat proteome does not show age-related susceptibility to oxidative damage nor increased ubiquitination [73]. Traditionally, the main mammalian models used for the study of ageing have been mice and rats, i.e. short-lived species [70]. To reach a reliable generalization and to deal with life-extending mechanisms, a long-lived species such as *H. glaber* can be precious; accordingly, in the past two decades, long-lived mammals (i.e. mole-rats and bats) became a matter of investigation. These new models drew attention to the oversimplification problem affecting existing theories about ageing in Chiroptera, and to several features considered universal components of enhanced longevity in mammals [70]. Indeed, among mammals with high longevity considerable variation exists with respect to candidate regulatory mechanisms [70].

Mitochondrial morphology and abundance were investigated in naked mole-rats, leading to interesting findings. In

contrast to mitochondria of other mammalian cardiomyocytes, where the internal space is filled with parallel rows of cristae, in naked mole-rats they have a chaotic pattern of wave-like cristae. In addition, there are some large mitochondria exceeding normal sizes by two to three times, with cristae assembled into groups with curved and ring-like structures [68]. Moreover, in these animals, mitochondrial number and size increase with age. They also possess a neotenic type of chondriome accompanied by specific features of OXPHOS, and a significant decrease in the level of matrix adenine nucleotides [71].

RNA sequencing was used to compare liver gene expression profiles between naked mole-rats and wild-derived mice. Genes associated with redox processes and mitochondria were expressed at higher relative levels in naked mole-rats. Also, a protease inhibitor and a mitochondrial complex II subunit, both ageing-related genes, were found strongly over-expressed in the naked mole-rat, suggesting alterations in mitochondrial and oxidation-reduction pathways in the species [74]. A combination of bioinformatic algorithms with nucleotide genomic signal processing and hierarchical cluster methods were also used to investigate differential phylogenetically convergent genetic traits related to senescence of the long-lived naked mole-rat and various species. Two ageing-related mitochondrial genes were selected based on the classification of ageing-related genes in the Human Ageing Genomic Resources database. Cytochrome *b* and cytochrome *c* oxidase subunit I are both related to neurodegenerative disease [72].

(d) Bats (*Myotis* spp.): heteroplasmy, longevity and ageing

The 'free radical theory of ageing' [109] posits that a high metabolic rate causes mitochondrial heteroplasmy and progressive ageing (but see [103]).The extended longevity of bats, despite their high metabolic rate (they are the only mammals with the ability of powered flight), may provide insights into mechanisms of ageing. Predictions of the free-radical theory of ageing as an explanation for differences in lifespan were tested in *Myotis lucifugus* (little brown bat), *Blarina brevicauda* (northern short-tailed shrew) and *P. leucopus* (white-footed mouse)—maximum lifespan potential: 34, 2, 8 years, respectively—by comparing whole-organism oxygen consumption, hydrogen peroxide production and superoxide dismutase activity in several tissues. Mitochondria from *M. lucifugus* produced half to one-third the amount of hydrogen peroxide per unit of oxygen consumed compared to mitochondria from *B. brevicauda* and *P. leucopus*, respectively. These results are similar to those reported for birds that, like bats, show high metabolic rates and longevity. These results seem to provide support, at least partially, for the free radical theory of ageing as an explanation to longevity of bats [76]. A series of genome-wide comparative analyses between bat (*Myotis myotis*) and non-bat mammals was performed to investigate the molecular mechanisms underlying longevity. Bat-specific and differentially transcribed micro RNA and messenger RNA that function in key longevity pathways were detected. As suggested for *M. myotis*, bats may possess unique regulatory mechanisms for resisting tumorigenesis, repairing cellular damage and preventing oxidative stress, all of which probably contribute to their extraordinary lifespan [78].

In this connection, analyses were done to see if bats exhibit increased mitochondrial heteroplasmy with age. In *M. myotis*, the majority of heteroplasmy was at a low frequency, and nucleotide changes consisted of transitions. Oxidative mutations were present in only a small number of individuals, there was no significant increase in heteroplasmy with age, and heteroplasmy was found to be dynamic (from recaptured individuals), without a uniform increase over time. All that considered, the data acquired for *M. myotis* appeared not to follow the predicted increase in heteroplasmy as posited by the free-radical theory of ageing, questioning its general validity [79]. On average, bats and birds live substantially longer than non-flying mammals of similar body size. The combination of small body size, high metabolic rates and long lifespan they share does not seem to support oxidative theories of ageing. However, large-scale comparative analyses on a few emerging model species have identified several mechanisms for resisting oxidative damage in mtDNA and cellular structures in both bats and birds [80].

(e) Eastern yellow robin (*Eopsaltria australis*): environmental adaptation, mitonuclear coevolution

Diversifying selection on metabolic pathways can reduce intraspecific gene flow and promote population divergence. An opportunity to explore this arises from mitonuclear discordance observed in an Australian bird *Eopsaltria australis*. Across 1500+ km, nuclear differentiation is low, whereas two highly divergent, parapatric mitochondrial lineages show a discordant longitudinal geographical pattern and experience different climates. *Eopsaltria australis* is one of several animals that show functional mitonuclear interactions despite discordance between mitochondrial and nuclear genomes. The two mitolineages show sharp climate-correlated differences in their distributions, suggesting that the mitochondrial introgression and divergence were driven by natural selection [85]. Evidence of positive selection has been found in genes of OXPHOS complexes, predicted to cause differences in electrostatic subunit-subunit interactions influencing coupling efficiency of the complex [83,85]. This selection may reflect local environmental adaptation, a by-product of other selective processes, or genetic incompatibilities [83].

Comparison of fixation indexes in the nuclear genomes between Eastern yellow robin populations across their biogeographic range revealed the existence of two genomic islands of divergence against a background of low differentiation [85]. A strong genetic differentiation and sequence divergence was detected in a region of chromosome 1A mirroring the geographical pattern of mtDNA divergence. Such a region is enriched for genes performing mitochondrial functions. Molecular signatures of selective sweeps in this region alongside those in the mitochondrial genome suggest a history of adaptive mitonuclear co-introgression, consistent with mitonuclear coevolution as an important mechanism for population divergence and local adaptation [84]. Studies on a large set of Australian birds were performed to investigate whether Pleistocene climate changes drove mitochondrial selection and evolution. Climate was a significant predictor of mitochondrial variation in eight out of 17 species [82].

(f) *Tigriopus californicus*: environmental adaptation, mitonuclear coevolution

The copepod *Tigriopus californicus* shows a very rapid mtDNA sequence evolution [94], reaching extreme mtDNA divergence

between populations with exceptionally high levels of amino acid differentiation compared with other taxa [89,88]. Hybridization between allopatric populations expose functional incompatibilities between genomes that have not coevolved and places intrinsic selective pressures on many nuclear genes involved in mitochondrial functions [86]. Ellison & Burton [90] measured fitness, mitochondrial function and enzyme activity in hybrid lines of *T. californicus*, and showed that only the OXPHOS complexes composed of both nuclear and mitochondrial subunits suffered a loss of activity. Moreover, they observed a positive correlation between OXPHOS enzyme activity and mitochondrial function, and between mitochondrial function and fitness. Moreover, the analysis of transcript levels of genes involved in the OXPHOS pathway for a series of parental and inbred hybrid lines showed that only genes located on the mtDNA differed among lines. Lines bearing certain genotypic combinations of mtDNA-mitochondrial RNA polymerase showed a diminished capacity to upregulate mitochondrial genes in response to hypo-osmotic stress. This suggests that disruption of coadaptation between nuclear and mitochondrial genes contributes to the phenomenon of hybrid breakdown [90,91,110]. *Tigriopus californicus* nuclear genome shows elevated protein evolutionary rates and putative positive selection in genes predicted to interact with mtDNA, mtRNAs and the corresponding proteins encoded. Thus, the rapid mitochondrial evolution appears to drive compensatory nuclear evolution within populations, thereby providing a potentially important mechanism for reproductive isolation [87].

As reported above for certain non-model vertebrates, variation in thermal tolerance plays a key role in determining the biogeographic distribution of organisms, and variation in mitochondrial function probably contributes to variation in thermal tolerance. Populations of *T. californicus* were also studied in this concern, showing a latitudinal thermal gradient along the coast of California (USA), suggesting thermal adaptation. The obtained data supported population-specific rates of ATP synthesis at chronic temperatures with a role for mitochondria in setting thermal range limits, and indicate that divergence in mitochondrial function is probably a component of adaptation across latitudinal thermal gradients [93].

Tigriopus californicus differs from standard animal models used in genetics because it lacks heteromorphic chromosomes but has a polygenic sex-determination mode. Reciprocal F_2 hybrids between two strongly differentiated populations were used for quantitative trait loci mapping. Although many studies on standard speciation models found the strongest genetic incompatibilities to be nuclear-nuclear (i.e.: X chromosome-autosome), in *T. californicus* the strongest deleterious interaction was mito-nuclear. This system thus is important to study the genetics of reproductive isolation in comparison to standard model systems [92].

(g) Bivalve molluscs: heteroplasmy, mitochondrial inheritance, mitonuclear interactions and genomic conflicts

More than a hundred species of bivalve molluscs are known for their unique mode of mitochondrial transmission, known since 1994 as doubly uniparental inheritance (DUI) (see [107] for a review). DUI species present two mitochondrial lineages, one transmitted by females (F-type), and one by males (M-type), this is achieved by having homoplasmic gametes: eggs for the F-type, spermatozoa for the M-type. Adult animals are heteroplasmic to various degrees, with interspecific mtDNA nucleotide divergence reaching exceptionally high values ([101] and references therein). The unusual mechanism of inheritance and the natural heteroplasmy provide an interesting point of view about mitochondrial biology in general. For example, the 'division of labour hypothesis' [111] posits that because mtDNA is maternally inherited, female gametes would prevent oxidative stress and mtDNA damage by repressing OXPHOS, thus being quiescent genetic templates. Interestingly, DUI species have swimming spermatozoa transmitting their mtDNA for hundreds of million years, with no apparent consequence for species survival [103,100]. The presence of stable and natural heteroplasmy of highly divergent mtDNAs makes DUI animals ideal systems to investigate genomic conflicts [104,105], and how an individual, a tissue, a cell, an organelle can deal with multiple mitochondrial variants [101]. So far, the DUI system was investigated mainly at the mtDNA level [107], and with RNA-Seq [98,99,97,102], even if more recently mitochondrial phenotypes of different kinds has been analysed. For example, differences between F- and M-types in OXPHOS activity—inferred by bioinformatics tools [106], or by direct test on oxygen consumption and bioenergetics [96,48]—were recently examined. Such divergent phenotypes may be produced by amino acid substitutions at specific sites of OXPHOS subunits differentiating the F-type from the M-type [106]. According to these new findings, the membrane potential may be involved in mitochondrial segregation into the germline [32], and/or with sperm-specific adaptations [99,96,48]. These studies evaluated mitochondrial functions associated with sex-linked mtDNAs in DUI species—among which *Arctica islandica*, the oldest animal known to date [95]—, compared DUI species with bivalve species showing the common strict maternal inheritance of mitochondria, and compared homoplasmic with heteroplasmic tissues. The results suggest that the inheritance of sex-linked mitochondria not only could better fulfil different energetic requirements of sperm and eggs, but could also be associated with success in facing heteroplasmy and ageing [96,48,99].

5. The comparative method in the '-omics' era: future perspectives

In the last 15 years, biology has changed radically. New technologies have reshaped the way data are collected and, most importantly, the amount of data that it is possible to obtain at an increasingly lower cost. Above all, high-throughput sequencing (HTS) granted virtually all laboratories around the world access to volumes of data unthinkable before. As a result, 'big data' became the foundation of most of the research in life sciences, and 'data mining' became one of the most used methods. While such change represents a breakthrough innovation that has greatly improved the potential of solving increasingly complex problems, it also has downsides. The deluge of data generated gave rise to what Sydney Brenner called 'factory science' in his 2007 Vanderbilt Discovery Lecture entitled 'The Next 100 Years of Biology' (video available at: http://hdl.handle.net/1803/1088), citing his words: 'I like to call it low-input, high-throughput, no-output biology'. This departure from hypothesis-driven research towards

model-independent data mining was prophetically discussed in two papers by John Allen in 2001 [112,113]. Allen used Karl Popper's metaphor of 'the bucket and the searchlight' [114] indicating two approaches to knowledge: the big-data approach, represented by the bucket, as opposed to the hypothesis-driven research, represented by the searchlight. It is true that HTS has made more systems accessible, but maybe it has unified the approach, extinguishing other methodological approaches in its wake. Are we moving from 'a few model systems' to 'a few model technologies', and if so, which bottleneck is more consequential? We do not wish to demonize HTS and new technologies which will probably be the pillars of new discoveries in the next few years. That said, we need to use such augmented technical power wisely.

HTS has surely made research on non-model species more accessible, but there are still limitations, and often the optimism about the revealing potential of genomics/transcriptomics data is excessive. We briefly list the major difficulties the researchers have to face when applying the comparative method.

(a) Sequencing disparity

New sequencing technologies were supposed to facilitate the study of organismal biodiversity, by making non-model species more accessible. This is not happening. David *et al.* [115] analysed species representation of nonhuman eukaryotes in the NCBI Sequence Read Archive, and found out that while species richness has been increasing, species evenness is decreasing. They reported a growing bias in favour of a minority of species: the top 1% most studied organisms represent a larger proportion of the experiments. In the future, it will be necessary to reverse such a trend, trying to explore diversity much more extensively. Initiatives such as the Global Invertebrate Genomics Alliance (GIGA) aims to promote standards that facilitate comparative approaches to genomics and collaborations across the scientific community [116]. It might sound ridiculous to split animals into vertebrates (including a subphylum of Chordata) and invertebrates (including all the other animal phyla), but given the disparity of genomics resources available—mostly vertebrates—it has a practical use. During the Biogenomics 2017 Conference in Washington DC the Earth Biogenome Project was announced, which aims to sequence, catalogue, and characterize the genomes of all eukaryotic biodiversity over a period of 10 years [117]. This project might seem overly ambitious, but 10 years is a long time for science and technology, and this impressive collaborative effort will probably reduce sequence disparity and enhance our ability to do comparative analyses.

(b) What to compare and how?

When choosing a model or extrapolating findings to other organisms, it is extremely important to know the phylogenetic context [5,17]. Not only knowing the phylogenetic relationships among organisms provides information about relatedness and divergence time, but it is fundamental to assess character evolution, polarity and homology: all essential data for the comparative method. For most of the animal taxa, there is a lack of phylogenetic framework—even for taxa closely related with model organisms—and this results in a problematic assessment of biodiversity and evolutionary relationships. Not being able to reconstruct the evolutionary history of the organisms undermines the comparative method.

Another fundamental resource we need when performing comparative analyses is gene annotation, which is the assignment of an identity and a function. Annotation works by translating the information available for model or well-known species, to the species subject of study [118]. Such translation is mostly based on sequence similarity—with all the obvious limitations—and nowadays, it is usually performed through fully automated procedures. The massive amount of genomics data generated and the consequent lack of human curation is causing a decrease in annotation quality [119]. Annotation is a difficult and time-consuming task, but unfortunately, the reward for such effort is almost non-existent: it would be difficult to publish a curated genome annotation in a relevant journal. In general, producing data are much more rewarding than curating them; in Sydney Brenner's words: 'We presently have huge amounts of data that are largely unassimilated. I have discovered that while there are rewards for collecting data and distributing data, there is little, if any, support for organizing it' [120, p. 9].

Strictly related with annotation, there is another big challenge—perhaps the biggest one—for comparative biologists, that is the concept of homology. Homology refers to the historical continuity of characters and it is a fundamental concept in biology, nevertheless, it is 'conceptually highly elusive' [121]. The main difficulty with homology is basically the topic of this Issue: the loose connection between genotype and phenotype. To make meaningful evolutionary studies, we need to compare genes that share ancestry and that retained the same function through evolutionary time. The so-called 'ortholog conjecture' [122] states that orthologues—homologous genes originated by a speciation event—retain the same function in the course of evolution, while paralogues—homologous genes originated by a duplication event—tend to evolve new functions. Currently, the typical procedure in comparative analyses is strictly based on the 'ortholog conjecture', and consists in comparing putatively orthologous sequences identified by sequence similarity [123]. As usual, the reality is much more complicated. The multidomain architecture of most eukaryotic proteins, gene duplication, recombination events, exon shuffling, gene fusion and fission are only some of the several processes that go under the spot-on definition of 'molecular tinkering' [124], which together with hybridization and introgression make the notion of homology at the molecular level an 'elusive concept' (for a thorough discussion on homology and related concepts, see [125]). A possible solution to this problem is to compare gene regulatory networks instead of individual genes. Such 'character identity networks' as defined by Günter Wagner [121] would enable more meaningful comparisons.

(c) The importance of 'basic research'

The representativeness of a model cannot be established in advance. Steel [126, p. 4] defined the 'extrapolator's circle', a paradox of model organisms and comparative research: 'establishing the suitability of the model would require already possessing detailed knowledge of the causal relationship in the target, in which case extrapolation would be unnecessary'. The choice of models is an iterative process: we make starting assumptions, test them against our knowledge, and then adjust as necessary, given our new discoveries [5]. The key aspect of a model organism is whether a community of scientists is working with it. Implementing the comparative method using as many models as possible basically means doing

'curiosity-driven' science (also known as basic research). Right now, basic research is underfunded, and mostly overlooked. We hope that, albeit more than 90 years later, an increasing number of biologists will start following the words of, among others, August Krogh and Hans Krebs. That would bring more balance between the currently predominant mechanistic/reductionist approach and the comparative approach.

Data accessibility. This article has no additional data.

Authors' contributions. F.G. and L.M. contributed equally to this work.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the Italian Ministry of Education, University and Research (MIUR) SIR Programme grant no. RBSI14G0P5 funded to L.M., MIUR FIR2013 Programme grant no. RBFR13T97A and the Canziani bequest funded to F.G., and 'Ricerca Fondamentale Orientata' (RFO) funding from the University of Bologna to L.M. and F.G.

Acknowledgements. We thank Helen Eaton, Senior Commissioning Editor, and the anonymous reviewers.

References

1. Krogh A. 1929 The progress of physiology. *Am. J. Physiol. Legacy Content* **90**, 243–251. (doi:10.1152/ajplegacy.1929.90.2.243)
2. Wang T. 2011 Gas exchange in frogs and turtles: how ectothermic vertebrates contributed to solving the controversy of pulmonary oxygen secretion. *Acta Physiol.* **202**, 593–600. (doi:10.1111/j.1748-1716.2011.02315.x)
3. Krebs HA. 1975 The August Krogh principle: 'for many problems there is an animal on which it can be most conveniently studied'. *J. Exp. Zool.* **194**, 221–226. (doi:10.1002/jez.1401940115)
4. Levy A, Currie A. 2015 Model organisms are not (theoretical) models. *Br. J. Phil. Sci.* **66**, 327–348. (doi:10.1093/bjps/axt055)
5. Bolker JA. 2009 Exemplary and surrogate models: two modes of representation in biology. *Perspect. Biol. Med.* **52**, 485–499. (doi:10.1353/pbm.0.0125)
6. Henderson L. 2019 *The problem of induction. The Stanford encyclopedia of philosophy*. Stanford, CA: The Metaphysics Research Laboratory.
7. Kabat GC. 2017 Taking distrust of science seriously: to overcome public distrust in science, scientists need to stop pretending that there is a scientific consensus on controversial issues when there is not. *EMBO Rep.* **18**, 1052–1055. (doi:10.15252/embr.201744294)
8. Sturtevant AH. 1959 Thomas Hunt Morgan. *Biograph. Mem. Natl Acad. Sci.* **33**, 283–325.
9. Crick F. 1970 Central dogma of molecular biology. *Nature* **227**, 561–563. (doi:10.1038/227561a0)
10. Brenner S, Sejnowski T. 2018 *In the spirit of science: lectures by Sydney Brenner on DNA*, worms and brains. Singapore: World Scientific.
11. Friedmann HC, Friedmann HC. 2004 From *Butyribacterium* to *E. coli*: an essay on unity in biochemistry. *Perspect. Biol. Med.* **47**, 47–66. (doi:10.1353/pbm.2004.0007)
12. Jørgensen CB. 2001 August Krogh and Claude Bernard on basic principles in experimental physiology. *Bioscience* **51**, 59–61. (doi:10.1641/0006-3568(2001)051[0059:AKACBO]2.0.CO;2)
13. Krebs HA, Krebs JR. 1980 The 'August Krogh principle'. *Comp. Biochem. Physiol. B: Comp. Biochem.* **67**, 379–380. (doi:10.1016/0305-0491(80)90324-7)
14. Gest H. 1995 Arabidopsis to Zebrafish: a commentary on Rosetta Stone model systems in the biological sciences. *Perspect. Biol. Med.* **39**, 77–85. (doi:10.1353/pbm.1995.0016)
15. Cohen SS. 1963 On biochemical variability and innovation. *Science* **139**, 1017–1026. (doi:10.1126/science.139.3559.1017)
16. Holmes FL. 1993 The old martyr of science: the frog in experimental physiology. *J. Hist. Biol.* **26**, 311–328. (doi:10.1007/BF01061972)
17. Kellogg EA, Shaffer HB. 1993 Model organisms in evolutionary studies. *Syst. Biol.* **42**, 409–414. (doi:10.1093/sysbio/42.4.409)
18. Green S, Dietrich MR, Leonelli S, Ankeny RA. 2018 'Extreme' organisms and the problem of generalization: interpreting the Krogh principle. *Hist. Phil. Life Sci.* **40**, 65. (doi:10.1007/s40656-018-0231-0)
19. Jacob F. 2001 Complexity and tinkering. *Ann. NY Acad. Sci.* **929**, 71–73. (doi:10.1111/j.1749-6632.2001.tb05708.x)
20. Wagner A. 2017 Information theory, evolutionary innovations and evolvability. *Phil. Trans. R. Soc. B* **372**, 20160416. (doi:10.1098/rstb.2016.0416)
21. Hunter P. 2008 The paradox of model organisms. The use of model organisms in research will continue despite their shortcomings. *EMBO Rep.* **9**, 717–720. (doi:10.1038/embor.2008.142)
22. Adriaens D, Herrel A. 2009 Functional consequences of extreme morphologies in the craniate trophic system. *Physiol. Biochem. Zool.* **82**, 1–6. (doi:10.1086/594382)
23. Münsterberg H. 1899 Psychology and history. *Psychol. Rev.* **6**, 1–3. (doi:10.1037/h0071306)
24. Hurlburt RT, Knapp TJ. 2006 Münsterberg in 1898, Not Allport in 1937, introduced the terms 'Idiographic' and 'Nomothetic' to American psychology. *Theory Psychol.* **16**, 287–293. (doi:10.1177/0959354306062541)
25. Jenner RA. 2008 Evo-devo's identity: from model organisms to developmental types. In *Evolving pathways – key themes in evolutionary developmental biology* (eds A Minelli, G Fusco), pp. 100–119. Cambridge, UK: Cambridge University Press.
26. Gould SJ. 1980 The promise of paleobiology as a nomothetic, evolutionary discipline. *Paleobiology* **6**, 96–118. (doi:10.1017/S0094837300012537)
27. Ghiselin MT. 2005 Homology as a relation of correspondence between parts of individuals. *Theory Biosci.* **124**, 91–103. (doi:10.1007/BF02814478)
28. Bolker JA, Raff RA. 1997 Beyond worms, flies, and mice: it's time to widen the scope of developmental biology. *J. NIH Res.* **9**, 35–39.
29. Scholtz G. 2008 On comparisons and causes in evolutionary developmental biology. In *Evolving pathways – key themes in evolutionary developmental biology* (eds A Minelli, G Fusco), pp. 144–159. Cambridge, UK: Cambridge University Press.
30. Russell JJ *et al.* 2017 Non-model model organisms. *BMC Biol.* **15**, 55. (doi:10.1186/s12915-017-0391-5)
31. Extavour CGM. 2008 Urbisexuality: the evolution of bilaterian germ cell specification and reproductive systems. In *Evolving pathways: key themes in evolutionary developmental biology* (eds A Minelli, G Fusco), pp. 321–342. Cambridge, UK: Cambridge University Press.
32. Milani L. 2015 Mitochondrial membrane potential: a trait involved in organelle inheritance? *Biol. Lett.* **11**, 20150732. (doi:10.1098/rsbl.2015.0732)
33. Isaeva VV, Reunov AA. 2001 Germ plasm and germline cell determination: the role of mitochondria. *Russ. J. Mar. Biol.* **27**, S8–S14. (doi:10.1023/A:1013842820458)
34. Milani L, Maurizii MG. 2019 Insights into germline development and differentiation in molluscs and reptiles: the use of molecular markers in the study of non-model animals. *Results Probl. Cell Differ.* **68**, 321–353. (doi:10.1007/978-3-030-23459-1_14)
35. Reunov A, Alexandrova Y, Reunova Y, Komkova A, Milani L. 2019 Germ plasm provides clues on meiosis: the concerted action of germ plasm granules and mitochondria in gametogenesis of the clam *Ruditapes philippinarum*. *Zygote* **27**, 25–35. (doi:10.1017/S0967199418000588)
36. Amikura R, Sato K, Kobayashi S. 2005 Role of mitochondrial ribosome-dependent translation in germline formation in *Drosophila* embryos. *Mech. Dev.* **122**, 1087–1093. (doi:10.1016/j.mod.2005.06.003)
37. Ballard JW, Whitlock MC. 2004 The incomplete natural history of mitochondria. *Mol. Ecol.* **13**, 729–744. (doi:10.1046/j.1365-294X.2003.02063.x)
38. Meiklejohn CD, Montooth KL, Rand DM. 2007 Positive and negative selection on the mitochondrial genome. *Trends Microbiol.* **23**, 259–263. (doi:10.1016/j.tig.2007.03.008)
39. Ballard JWO, Pichaud N. 2014 Mitochondrial DNA: more than an evolutionary bystander. *Funct. Ecol.* **28**, 218–231. (doi:10.1111/1365-2435.12177)
40. Dowling DK. 2014 Evolutionary perspectives on the links between mitochondrial genotype and disease phenotype. *Biochim. Biophys. Acta* **1840**, 1393–1403. (doi:10.1016/j.bbagen.2013.11.013)

41. Hill GE. 2015 Mitonuclear ecology. *Mol. Biol. Evol.* **32**, 1917–1927. (doi:10.1093/molbev/msv104)
42. Hill GE, Havird JC, Sloan DB, Burton RS, Greening C, Dowling DK. 2018 Assessing the fitness consequences of mitonuclear interactions in natural populations. *Biol. Rev.* **2**, 3064. (doi:10.1111/brv.12493)
43. Sloan DB, Warren JM, Williams AM, Wu Z, Abdel-Ghany SE, Chicco AJ, Havird JC. 2018 Cytonuclear integration and co-evolution. *Nat. Rev. Genet.* **19**, 635–648. (doi:10.1038/s41576-018-0035-9)
44. Havird JC, Forsythe ES, Williams AM, Werren JH, Dowling DK, Sloan DB. 2019 Selfish mitonuclear conflict. *Curr. Biol.* **29**, R496–R511. (doi:10.1016/j.cub.2019.03.020)
45. Havird JC *et al.* 2019 Beyond the powerhouse: integrating mitonuclear evolution, physiology, and theory in comparative biology. *Integr. Comp. Biol.* **59**, 856–863. (doi:10.1093/icb/icz132)
46. Hill GE. 2019 *Mitonuclear ecology*. Oxford, UK: Oxford University Press.
47. Breton S, Milani L, Ghiselli F, Guerra D, Stewart DT, Passamonti M. 2014 A resourceful genome: updating the functional repertoire and evolutionary role of animal mitochondrial DNAs. *Trends Genet.* **30**, 555–564. (doi:10.1016/j.tig.2014.09.002)
48. Bettinazzi S, Nadarajah S, Dalpé A, Milani L, Blier PU, Breton S. 2019 Linking paternally inherited mtDNA variants and sperm performance. *Phil. Trans. R. Soc. B* **375**, 20190177. (doi:10.1098/rstb.2019.0177)
49. Elbassiouny AA, Lovejoy NR, Chang BSW. 2019 Convergent patterns of evolution of mitochondrial oxidative phosphorylation (OXPHOS) genes in electric fishes. *Phil. Trans. R. Soc. B* **375**, 20190179. (doi:10.1098/rstb.2019.0179)
50. Havird JC, Shah AA, Chicco AJ. 2019 Powerhouses in the cold: mitochondrial function during thermal acclimation in montane mayflies. *Phil. Trans. R. Soc. B* **375**, 20190181. (doi:10.1098/rstb.2019.0181)
51. Baris TZ, Crawford DL, Oleksiak MF. 2016 Acclimation and acute temperature effects on population differences in oxidative phosphorylation. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **310**, R185–R196. (doi:10.1152/ajpregu.00421.2015)
52. Chung DJ, Bryant HJ, Schulte PM. 2017 Thermal acclimation and subspecies-specific effects on heart and brain mitochondrial performance in a eurythermal teleost (*Fundulus heteroclitus*). *J. Exp. Biol.* **220**, 1459–1471. (doi:10.1242/jeb.151217)
53. Chung DJ, Morrison PR, Bryant HJ, Jung E, Brauner CJ, Schulte PM. 2017 Intraspecific variation and plasticity in mitochondrial oxygen binding affinity as a response to environmental temperature. *Sci. Rep.* **7**, 16238. (doi:10.1038/s41598-017-16598-6)
54. Chung DJ, Healy TM, McKenzie JL, Chicco AJ, Sparagna GC, Schulte PM. 2018 Mitochondria, temperature, and the pace of life. *Integr. Comp. Biol.* **58**, 578–590. (doi:10.1093/icb/icy013)
55. Fangue NA, Richards JG, Schulte PM. 2009 Do mitochondrial properties explain intraspecific variation in thermal tolerance? *J. Exp. Biol.* **212**, 514–522. (doi:10.1242/jeb.024034)
56. McKenzie JL, Dhillon RS, Schulte PM. 2016 Steep, coincident, and concordant clines in mitochondrial and nuclear-encoded genes in a hybrid zone between subspecies of Atlantic killifish, *Fundulus heteroclitus*. *Ecol. Evol.* **6**, 5771–5787. (doi:10.1002/ece3.2324)
57. Whitehead A, Roach JL, Zhang S, Galvez F. 2012 Salinity- and population-dependent genome regulatory response during osmotic acclimation in the killifish (*Fundulus heteroclitus*) gill. *J. Exp. Biol.* **215**, 1293–1305. (doi:10.1242/jeb.062075)
58. Baumgart M *et al.* 2016 Longitudinal RNA-Seq analysis of vertebrate aging identifies mitochondrial complex i as a small-molecule-sensitive modifier of lifespan. *Cell Syst.* **2**, 122–132. (doi:10.1016/j.cels.2016.01.014)
59. Harel I *et al.* 2015 A platform for rapid exploration of aging and diseases in a naturally short-lived vertebrate. *Cell* **160**, 1013–1026. (doi:10.1016/j.cell.2015.01.038)
60. Hartmann N *et al.* 2011 Mitochondrial DNA copy number and function decrease with age in the short-lived fish *Nothobranchius furzeri*. *Aging Cell* **10**, 824–831. (doi:10.1111/j.1474-9726.2011.00723.x)
61. Cheviron ZA, Bachman GC, Connaty AD, McClelland GB, Storz JF. 2012 Regulatory changes contribute to the adaptive enhancement of thermogenic capacity in high-altitude deer mice. *Proc. Natl Acad. Sci. USA* **109**, 8635–8640. (doi:10.1073/pnas.1120523109)
62. Cheviron ZA, Connaty AD, McClelland GB, Storz JF. 2014 Functional genomics of adaptation to hypoxic cold-stress in high-altitude deer mice: transcriptomic plasticity and thermogenic performance. *Evolution* **68**, 48–62. (doi:10.1111/evo.12257)
63. Lui MA, Mahalingam S, Patel P, Connaty AD, Ivy CM, Cheviron ZA, Storz JF, McClelland GB, Scott GR. 2015 High-altitude ancestry and hypoxia acclimation have distinct effects on exercise capacity and muscle phenotype in deer mice. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **308**, R779–R791. (doi:10.1152/ajpregu.00362.2014)
64. Mahalingam S, McClelland GB, Scott GR. 2017 Evolved changes in the intracellular distribution and physiology of muscle mitochondria in high-altitude native deer mice. *J. Physiol.* **595**, 4785–4801. (doi:10.1113/JP274130)
65. Nikel KE, Shanishchara NK, Ivy CM, Dawson NJ, Scott GR. 2018 Effects of hypoxia at different life stages on locomotory muscle phenotype in deer mice native to high altitudes. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **224**, 98–104. (doi:10.1016/j.cbpb.2017.11.009)
66. Scott GR, Elogio TS, Lui MA, Storz JF, Cheviron ZA. 2015 Adaptive modifications of muscle phenotype in high-altitude deer mice are associated with evolved changes in gene regulation. *Mol. Biol. Evol.* **32**, 1962–1976. (doi:10.1093/molbev/msv076)
67. Scott GR, Guo KH, Dawson NJ. 2018 The mitochondrial basis for adaptive variation in aerobic performance in high-altitude deer mice. *Integr. Comp. Biol.* **58**, 506–518. (doi:10.1093/icb/icy056)
68. Bakeeva L, Vays V, Vangeli I, Eldarov C, Holtze S, Hildebrandt T, Skulachev V. 2019 Delayed onset of age-dependent changes in ultrastructure of myocardial mitochondria as one of the neotenic features in naked mole rats (*Heterocephalus glaber*). *Int. J. Mol. Sci.* **20**, piiE566. (doi:10.3390/ijms20030566)
69. Buffenstein R. 2005 The naked mole-rat: a new long-living model for human aging research. *J. Gerontol. A Biol. Sci. Med. Sci.* **60**, 1369–1377. (doi:10.1093/gerona/60.11.1369)
70. Dammann P. 2017 Slow aging in mammals: lessons from African mole-rats and bats. *Semin. Cell Dev. Biol.* **70**, 154–163. (doi:10.1016/j.semcdb.2017.07.006)
71. Holtze S, Eldarov CM, Vays VB, Vangeli IM, Vysokikh MY, Bakeeva LE, Skulachev VP, Hildebrandt TB. 2016 Study of age-dependent structural and functional changes of mitochondria in skeletal muscles and heart of naked mole rats (*Heterocephalus glaber*). *Biochemistry* **81**, 1429–1437. (doi:10.1134/s000629791612004x)
72. Khayal LA, Provaznik I, Tkacz E. 2013 Differential analysis of neurodegenerative aging-related mitochondrial genes of long-lived naked mole-rat. *Int. J. Biosci. Biochem. Bioinforma* **3**, 75. (doi:10.7763/ijbbb.2013.v3.168)
73. Kim EB *et al.* 2011 Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature* **479**, 223–227. (doi:10.1038/nature10533)
74. Yu C *et al.* 2011 RNA sequencing reveals differential expression of mitochondrial and oxidation reduction genes in the long-lived naked mole-rat when compared to mice. *PLoS ONE* **6**, e26729. (doi:10.1371/journal.pone.0026729)
75. Ball HC, Levari-Shariati S, Cooper LN, Aliani M. 2018 Comparative metabolomics of aging in a long-lived bat: insights into the physiology of extreme longevity. *PLoS ONE* **13**, e0196154. (doi:10.1371/journal.pone.0196154)
76. Brunet-Rossinni AK. 2004 Reduced free-radical production and extreme longevity in the little brown bat (*Myotis lucifugus*) versus two non-flying mammals. *Mech. Ageing Dev.* **125**, 11–20. (doi:10.1016/j.mad.2003.09.003)
77. Brunet-Rossinni AK, Wilkinson GS. 2009 Methods for age estimation and the study of senescence in bats. In *Ecological and behavioral methods for the study of bats* (eds TH Kunz, S Parsons), pp. 315–325. Johns Hopkins University Press.
78. Huang Z, Jebb D, Teeling EC. 2016 Blood miRNomes and transcriptomes reveal novel longevity mechanisms in the long-lived bat, *Myotis myotis*. *BMC Genomics* **17**, 906. (doi:10.1186/s12864-016-3227-8)
79. Jebb D, Foley NM, Whelan CV, Touzalin F, Puechmaille SJ, Teeling EC. 2018 Population level mitogenomics of long-lived bats reveals dynamic heteroplasmy and challenges the free radical theory of ageing. *Sci. Rep.* **8**, 13634. (doi:10.1038/s41598-018-31093-2)
80. Munshi-South J, Wilkinson GS. 2010 Bats and birds: exceptional longevity despite high metabolic rates. *Ageing Res. Rev.* **9**, 12–19. (doi:10.1016/j.arr.2009.07.006)
81. Seim I *et al.* 2013 Genome analysis reveals insights into physiology and longevity of the Brandt's bat *Myotis brandtii*. *Nat. Commun.* **4**, 2212. (doi:10.1038/ncomms3212)
82. Lamb AM, Gan HM, Greening C, Joseph L, Lee YP, Morán-Ordóñez A, Sunnucks P, Pavlova A. 2018 Climate-driven mitochondrial selection: a test in Australian songbirds. *Mol. Ecol.* **27**, 898–918. (doi:10.1111/mec.14488)

83. Morales HE, Pavlova A, Joseph L, Sunnucks P. 2015 Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. *Mol. Ecol.* **24**, 2820–2837. (doi:10.1111/mec.13203)
84. Morales HE, Pavlova A, Amos N, Major R, Kilian A, Greening C, Sunnucks P. 2018 Concordant divergence of mitogenomes and a mitonuclear gene cluster in bird lineages inhabiting different climates. *Nat. Ecol. Evol.* **2**, 1258–1267. (doi:10.1038/s41559-018-0606-3)
85. Sunnucks P, Morales HE, Lamb AM, Pavlova A, Greening C. 2017 Integrative approaches for studying mitochondrial and nuclear genome co-evolution in oxidative phosphorylation. *Front. Genet.* **8**, 25. (doi:10.3389/fgene.2017.00025)
86. Barreto FS, Burton RS. 2013 Evidence for compensatory evolution of ribosomal proteins in response to rapid divergence of mitochondrial rRNA. *Mol. Biol. Evol.* **30**, 310–314. (doi:10.1093/molbev/mss228)
87. Barreto FS, Watson ET, Lima TG, Willett CS, Edmands S, Li W, Burton RS. 2018 Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nat. Ecol. Evol.* **2**, 1250–1257. (doi:10.1038/s41559-018-0588-1)
88. Burton RS, Byrne RJ, Rawson PD. 2007 Three divergent mitochondrial genomes from California populations of the copepod *Tigriopus californicus*. *Gene* **403**, 53–59. (doi:10.1016/j.gene.2007.07.026)
89. Edmands S. 2001 Phylogeography of the intertidal copepod *Tigriopus californicus* reveals substantially reduced population differentiation at northern latitudes. *Mol. Ecol.* **10**, 1743–1750. (doi:10.1046/j.0962-1083.2001.01306.x)
90. Ellison CK, Burton RS. 2006 Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus californicus*. *Evolution* **60**, 1382–1391. (doi:10.1111/j.0014-3820.2006.tb01217.x)
91. Ellison CK, Burton RS. 2008 Interpopulation hybrid breakdown maps to the mitochondrial genome. *Evolution* **62**, 631–638. (doi:10.1111/j.1558-5646.2007.00305.x)
92. Foley BR, Rose CG, Rundle DE, Leong W, Edmands S. 2013 Postzygotic isolation involves strong mitochondrial and sex-specific effects in *Tigriopus californicus*, a species lacking heteromorphic sex chromosomes. *Heredity* **111**, 391–401. (doi:10.1038/hdy.2013.61)
93. Harada AE, Healy TM, Burton RS. 2019 Variation in thermal tolerance and its relationship to mitochondrial function across populations of *Tigriopus californicus*. *Front. Physiol.* **10**, 213. (doi:10.3389/fphys.2019.00213)
94. Willett CS. 2012 Quantifying the elevation of mitochondrial DNA evolutionary substitution rates over nuclear rates in the intertidal copepod *Tigriopus californicus*. *J. Mol. Evol.* **74**, 310–318. (doi:10.1007/s00239-012-9508-1)
95. Butler PG, Wanamaker AD, Scourse JD, Richardson CA, Reynolds DJ. 2013 Variability of marine climate on the North Icelandic Shelf in a 1357-year proxy archive based on growth increments in the bivalve *Arctica islandica*. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **373**, 141–151. (doi:10.1016/j.palaeo.2012.01.016)
96. Bettinazzi S, Rodríguez E, Milani L, Blier PU, Breton S. 2019 Metabolic remodelling associated with mtDNA: insights into the adaptive value of doubly uniparental inheritance of mitochondria. *Proc. R. Soc. B* **286**, 20182708. (doi:10.1098/rspb.2018.2708)
97. Capt C, Renaut S, Ghiselli F, Milani L, Johnson NA, Sietman BE, Stewart DT, Breton S. 2018 Deciphering the link between doubly uniparental inheritance of mtDNA and sex determination in bivalves: clues from comparative transcriptomics. *Genome Biol. Evol.* **10**, 577–590. (doi:10.1093/gbe/evy019)
98. Ghiselli F, Milani L, Chang PL, Hedgecock D, Davis JP, Nuzhdin SV, Passamonti M. 2012 De novo assembly of the Manila clam *Ruditapes philippinarum* transcriptome provides new insights into expression bias, mitochondrial doubly uniparental inheritance and sex determination. *Mol. Biol. Evol.* **29**, 771–786. (doi:10.1093/molbev/msr248)
99. Ghiselli F, Milani L, Guerra D, Chang PL, Breton S, Nuzhdin SV, Passamonti M. 2013 Structure, transcription, and variability of metazoan mitochondrial genome: perspectives from an unusual mitochondrial inheritance system. *Genome Biol. Evol.* **5**, 1535–1554. (doi:10.1093/gbe/evt112)
100. Ghiselli F, Breton S, Milani L. 2018 Mitochondrial activity in gametes and uniparental inheritance: a comment on 'What can we infer about the origin of sex in early eukaryotes?' *Phil. Trans. R. Soc. B* **373**, 20170147. (doi:10.1098/rstb.2017.0147)
101. Ghiselli F *et al.* 2019 Natural heteroplasmy and mitochondrial inheritance in bivalve molluscs. *Integr. Comp. Biol.* **59**, 1016–1032. (doi:10.1093/icb/icz061)
102. Iannello M, Puccio G, Piccinini G, Passamonti M, Ghiselli F. 2019 The dynamics of mito-nuclear coevolution: a perspective from bivalve species with two different mechanisms of mitochondrial inheritance. *J. Zoolog. Syst. Evol. Res.* **369**, 20130438. (doi:10.1111/jzs.12271)
103. Milani L, Ghiselli F. 2015 Mitochondrial activity in gametes and transmission of viable mtDNA. *Biol. Direct* **10**, 22. (doi:10.1186/s13062-015-0057-6)
104. Milani L, Ghiselli F, Passamonti M. 2016 Mitochondrial selfish elements and the evolution of biological novelties. *Curr. Zool.* **62**, 687–697. (doi:10.1093/cz/zow044)
105. Milani L, Ghiselli F, Pecci A, Maurizii MG, Passamonti M. 2015 The expression of a novel mitochondrially-encoded gene in gonadic precursors may drive paternal inheritance of mitochondira. *PLoS ONE* **10**, e0137468. (doi:10.1371/journal.pone.0137468)
106. Skibinski DOF, Ghiselli F, Diz AP, Milani L, Mullins JGL. 2017 Structure-related differences between cytochrome oxidase I proteins in a stable heteroplasmic mitochondrial system. *Genome Biol. Evol.* **9**, 3265–3281. (doi:10.1093/gbe/evx235)
107. Zouros E. 2013 Biparental inheritance through uniparental transmission: the doubly uniparental inheritance (DUI) of mitochondrial DNA. *Evol. Biol.* **40**, 1–31. (doi:10.1007/s11692-012-9195-2)
108. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013 The hallmarks of aging. *Cell* **153**, 1194–1217. (doi:10.1016/j.cell.2013.05.039)
109. Harman D. 1956 Aging: a theory based on free radical and radiation chemistry. *J. Gerontol.* **11**, 298–300. (doi:10.1093/geronj/11.3.298)
110. Rand DM, Mossman JA. 2019 Mitonuclear conflict and cooperation govern the integration of genotypes, phenotypes and environments. *Phil. Trans. R. Soc. B* **375**, 20190188. (doi:10.1098/rstb.2019.0188)
111. Allen JF. 1996 Separate sexes and the mitochondrial theory of ageing. *J. Theor. Biol.* **180**, 135–140. (doi:10.1006/jtbi.1996.0089)
112. Allen JF. 2001 *In silico* veritas. Data-mining and automated discovery: the truth is in there. *EMBO Rep.* **2**, 542–544. (doi:10.1093/embo-reports/kve139)
113. Allen JF. 2001 Bioinformatics and discovery: induction beckons again. *Bioessays* **23**, 104–107. (doi:10.1002/1521-1878(200101)23:1<104::AID-BIES1013>3.0.CO;2-2)
114. Popper KR. 1975 *Objective knowledge: an evolutionary approach*. Oxford, UK: Oxford University Press.
115. David KT, Wilson AE, Halanych KM. 2019 Sequencing disparity in the genomic era. *Mol. Biol. Evol.* **36**, 1624–1627. (doi:10.1093/molbev/msz117)
116. Voolstra CR, GIGA Community of Scientists (COS), Wörheide G, Lopez JV. 2017 Advancing genomics through the global invertebrate genomics alliance (GIGA). *Invertebr. Syst.* **31**, 1–7. (doi:10.1071/IS16059)
117. Lewin HA *et al.* 2018 Earth BioGenome project: sequencing life for the future of life. *Proc. Natl Acad. Sci. USA* **115**, 4325–4333. (doi:10.1073/pnas.1720115115)
118. Primmer CRR, Papakostas S, Leder EHH, Davis MJJ, Ragan MAA. 2013 Annotated genes and nonannotated genomes: cross-species use of gene ontology in ecology and evolution research. *Mol. Ecol.* **22**, 3216–3241. (doi:10.1111/mec.12309)
119. Salzberg SL. 2019 Next-generation genome annotation: we still struggle to get it right. *Genome Biol.* **20**, 92. (doi:10.1186/s13059-019-1715-2)
120. Friedberg EC. 2008 Sydney Brenner. *Nat. Rev. Mol. Cell Biol.* **9**, 8–9. (doi:10.1038/nrm2320)
121. Wagner GP. 2007 The developmental genetics of homology. *Nat. Rev. Genet.* **8**, 473–479. (doi:10.1038/nrg2099)
122. Nehrt NL, Clark WT, Radivojac P, Hahn MW. 2011 Testing the ortholog conjecture with comparative functional genomic data from mammals. *PLoS Comput. Biol.* **7**, e1002073. (doi:10.1371/journal.pcbi.1002073)
123. Koonin EV. 2005 Orthologs, paralogs, and evolutionary genomics. *Annu. Rev. Genet.* **39**, 309–338. (doi:10.1146/annurev.genet.39.073003.114725)
124. Graur D. 2016 *Molecular and genome evolution*. Sunderland, MA: Sinauer Associates.
125. Wagner GP. 2014 *Homology, genes, and evolutionary innovation*. Princeton, NJ: Princeton University Press.
126. Steel D. 2008 *Across the boundaries: extrapolation in biology and social science*. Oxford, UK: Oxford University Press.