PERSPECTIVE

National Science Open
2: 20220054, 2023
https://doi.org/10.1360/nso/20220054

Life Sciences and Medicine

Exploring mysteries of life and their great values in biomedicine through evolutionary genomics

Wen Wang[1,2,*]

[1]*School of Ecology and Environment, Northwestern Polytechnical University, Xi'an 710072, China;*
[2]*Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650201, China*

*Corresponding author (email: wenwang@nwpu.edu.cn)

Received 22 June 2022; Revised 13 December 2022; Accepted 23 December 2022; Published online 6 February 2023

Among the endless life forms evolved in nature, some have unusual capacity in a certain function compared to other species. The underlying molecular mechanism has been largely mysterious due to long-standing gap between microevolutionary and macroevolutionary studies as well as focus on model organisms in mainstream biological studies. In this article, I first discuss the conceptual evolution of evolutionary theories, and then use recent practical examples to show that the omics + across taxa + traits approach, which I refer to as "Evolutionary genomics" in a narrow sense, or "Grand synthesis" in a broad sense, can reveal genomics and molecular scenarios of miraculous mysteries of life, especially those having biomedical implications. It is expected that such evolutionary genomics studies across taxa will emerge in large numbers in the coming years, and may inspire more and more biomedical applications.

In the epoch-making work, *Origin of Species*, Charles Darwin [1] was amazed that "from so simple a beginning (a few forms or one) endless forms most beautiful and most wonderful have been, and are being, evolved". After billion years of evolution, endless beautiful organismal forms have evolved, and are still evolving. Many of these forms are superior to human beings in terms of adapting to harsh environments or capacity. For example, bats and cetaceans can emit ultrasonic wave to detect prey, and many insects and even some mammals can see ultraviolet or infrared light [2] (in the article Burnett described more unusual sensory powers of animals). Recently, a number of bionic applications of superior biological perception or structures have been reported, such as high-speed imaging of infrared light inspired by the iridescent Morpho butterfly scales [3], near-infrared light sensitivity in mouse models through transgene or nanoparticles [4,5], and biomimics of gecko spatulae [6].

However, animal traits relevant to human health may attract more attention in biological studies. In animals, there are many species with long lifespan compared to other species in the same taxa. One jellyfish species (*Turritopsis dohrnii*) is a kind of immortal organisms, which can resume to the juvenile stage from a mature adult stage again and again [7]. Even in mammals, bowhead whales can live as long as over 200 years, which is the mammalian species with the known longest lifespan [8]. More other animal traits with biomedical implications, such as low cancer rate, adaptation to hypoxia or polar environment, and regeneration capacity, have also been studied in recent years, some of which will be described in the following

text. Nevertheless, during the previous era of molecular biology, few model organisms have been the foci of biological studies, and thus endless beautiful forms in nature have largely been ignored for more than half a century. Beyond ignorance, in a sense, the paradigm of modern molecular biology has strong characteristics of reductionism and analyticalism, and thus it is difficult to deal with diverse and complex natural species. On the contrary, evolutionary biology, which was stemmed from natural history before Darwin's period, has always been trying to understand the pattern and mechanisms of biodiversity [9].

Lamarck was among the first pioneers who proposed organisms were from evolution rather than creation by the God [10]. Darwin then rejected Lamarckism's theory of acquired inheritance and identified natural selection as the driving force of the origin of species or evolution of organisms [1] (Figure 1A). In the early 20th century, after the rediscovery of Mendelian genetics, Haldane, Wright and Fisher integrated Mendelian genetics with Darwinian evolution and created the discipline "population genetics", and then by extending this, more evolutionary biologists, including Dobzhansky (1937) [11], Mayr (1942) [12] and Stebbins (1937) [13], coined the term of Modern Synthesis (MS) of evolutionary theory. Since then MS had become the dogmatic paradigm of evolutionary biology. It is a complex theoretical system and one sentence briefing of this theory could be "evolution is changes in allele frequencies within populations". MS still recognizes natural selection as the driving force of adaptive evolution although genetic mutations could be random as pointed out by the Neutral Theory [14]. For a time, the Neutral Theory of molecular evolution was by some taken to invalidate the Darwinian selection theory, but eventually it has become the conceptual null hypothesis to detect selection signals at the molecular level. The great success of MS has made the traditional evolutionary studies concentrate on allele frequency changes in populations, although in practice nowadays many evolutionary biologists have affiliated back to Darwin's original definition of evolution being "descent with modification". Unfortunately, before the genomics era, due to historical constraint especially immaturity and lack of necessary technology and research methods, traditional evolutionary studies were dominated by theoretical modeling of gene frequency and population genetics survey within species, resulting in great gaps between molecular genetic mutations and phenotypic innovations across taxa. Typical evolutionary genetics studies were barely able to link gene frequency (genotype) with origin of traits (phenotype), especially those big trait innovations in a large group of taxa, such as the water-to-land transition of vertebrates and origin of a new organ. Studies on big trait innovations have mainly been the domain of paleontology, or the so-called macroevolution. Therefore, evolutionary studies have been largely isolated between microevolution and macroevolution.

In order to bridge the gap between microevolution and macroevolution, many efforts have been tried in recent decades. In 1975, King and Wilson [15] published an influential paper which proposed the importance of regulatory mutations in phenotypic innovation because "the macromolecules between humans and chimpanzees are so alike that regulatory mutations may account for their biological differences". This idea greatly gestated the proposition of evolutionary developmental biology (for short, usually referred to as evo-devo) initiated by the discovery of homeobox genes, which was even called the second synthesis integrating embryology as well as molecular genetics, phylogeny, and evolutionary biology [16,17]. The core idea of evo-devo is that all phyla of animals share a few homeobox toolkit genes and different body plans result from different regulation of these genes. However, the inherent analyticalism feature of evo-devo and methodological dependence on few homeobox genes let it face insurmountable challenges to explain how endless animal forms have evolved. Trying to adopt an integrative approach to understand the phenotype-genotype

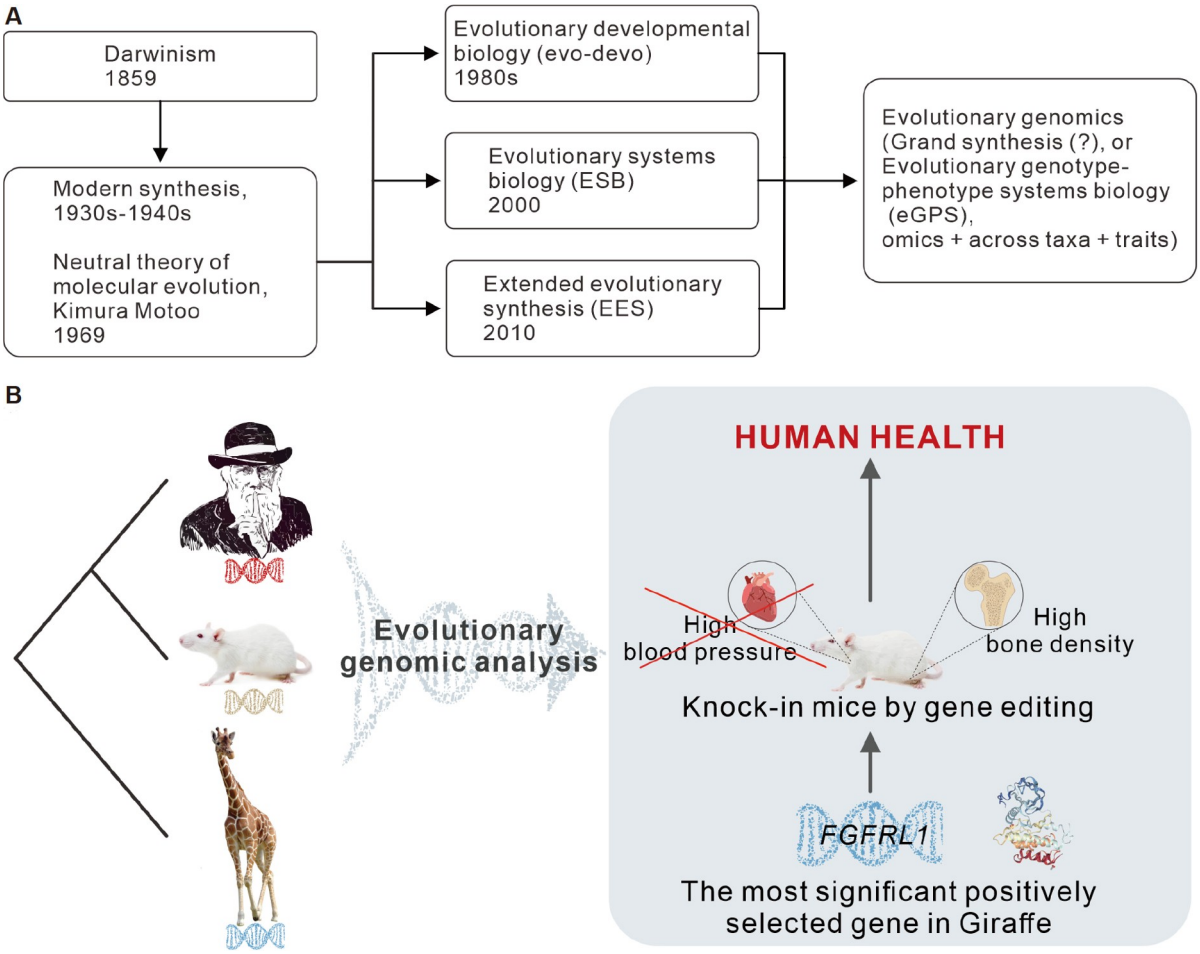

Figure 1 A graphic depiction of the evolution of evolutionary theories (A) and an example of recent evolutionary genomics study in which a positively selected gene found in giraffes can result in high bone density and resistance to hypertension (B) [32]. The question mark indicates that I am not sure that if the "Grand synthesis" is a proper name or not for the current evolutionary biology. The case in (B) vividly shows that the evolutionary genomics approach proposed in this paper can have great power to reveal molecular mechanism underpinning an important trait and enlighten future biomedical research in hypertension and osteoporosis [32].

relationships, a decade ago a new evolutionary discipline—evolutionary systems biology (ESB) was proposed [18,19]. However, still there has been no precise definition for ESB yet, and being represented by the characteristics of systems biology, it attempts to include all kinds of methodologies such as omics, networks, dynamic modeling and even synthetic biology. The currently blurred scheme and complex parameters of ESB result in a problem that it is hard to come up with a clear roadmap when one wants to tackle an evolutionary question in the context of ESB; for example, how had water-to-land transition been achieved? Conceptually, it is also not very clear what are the hallmarks of ESB except the idea of integrative and systems thinking.

A recent movement advocates an extended evolutionary synthesis (EES) to reflect several important conceptual modifications on the traditional MS [20,21]. The EES intriguingly pointed out that several phenomena require extension of the MS. First, developmental bias, i.e., the inherent genetic composition of ancestors, can restrain the diversification direction even in different environments, rather than convergent

evolution assumed by the traditional MS theory. Second, niche construction by organisms can modify environment and is thus beyond the conjecture of MS that organisms passively adapt to environmental changes. The most remarkable conceptual revision may be acknowledging the existence of acquired inheritances through trans-generation epigenetic changes [22]. Some evolutionary biologists think these term extensions are not necessary or even meaningless because all the phenomena have been or could be considered in the context of MS [20], but I think as least the acquired inheritance through epigenetics has not been an element of MS. In my opinion, to include many newest discoveries unexpected by MS and expand the capacity of bridging phenotype and genotype as well as microevolution and macroevolution are of great theoretical and methodological values, which will perhaps help resume evolutionary biology to the mainstream of biology.

Recently, we proposed an approach of evolutionary genotype-phenotype systems biology (eGPS) [23], aiming to integrate large-scale omics data across species and especially high taxonomic categories such as class and order level to reveal molecular and genetic mechanisms underpinning big and complex phenotypic innovations. The abbreviation form (eGPS) of this approach sounds echoing the GPS navigation system in the ability of localizing causal genetic mutations of phenotypes, but the full name is kind of awkward and too close to the previous ESB, and did not reflect the core idea: omics + across taxa + traits. Here I would like to refer to this paradigm as "Evolutionary Genomics" in a narrow and specific sense, or more ambitiously but less clearly "Grand Synthesis". The former may be more acceptable by the evolutionary biology community and its instruction to methodology is relatively clear, namely, omics + across taxa + traits, although the aim to reveal big trait innovation is still not that obvious. The narrow sense of Evolutionary Genomis is also very close, or even equivalent to the widely-used term of comparative genomics. The "Grand Synthesis" may be able to concisely include the purpose of ESB and EES to conceptually promote evolutionary study paradigm from reductionism to more integrative and systems ways, but the practical methodology is as relatively obscure as EES. Therefore, the proper name for the new paradigm or approach still awaits to be scrutinized by the community while the successes that using large omics data across large-scale evolutionary taxa to reveal genetic and molecular mechanisms underlying big trait innovations have been witnessed in the past few years.

A graphic depiction of the evolution of evolutionary theories described in this section is shown in Figure 1A.

The first most remarkable evolutionary genomics study could be the Bird 10000 Genome (B10K) project which used large-scale genome sequence information covering species of all orders of the bird class. This project resolved the phylogeny of Avian orders resulting from rapid radiation about 100 million years ago (Mya) [24]. Based on the well-revolved phylogeny, lineage-specific functional elements related to lifestyles and traits including vocal learning, flight, absence of teeth, and so on [25–28] were able to be inferred. These functional elements include positively selected genes, fast-evolving genes, and lineage-specific and conserved regulatory sequences, which would not be identified without well-resolved phylogeny and well-aligned syntenic genomic regions. The disclosure of genetic architectures of many avian traits displayed an unprecedented power in deciphering mysteries of life or species diversity through evolutionary genomics approach.

The avian case tackled a lineage which appeared 100 Mya, and structure of bird genomes is unusually conserved and compact compared to various other animals, a feature making the evolutionary genomics

study more feasible. For other taxa diverged hundreds of million years ago which were often accompanied with big phenotypic innovation events, the genomes to be compared could be so complex and diverged that regular comparative genomic and evolutionary genetic analyses are not applicable. In a recent effort we managed to decipher the ever-reported largest genome of the African lungfish (40 Gb), and compared it with the basal ray-finned fishes and tetrapods to tackle the question of water-to-land transition of vertebrates, which happened 480 Mya [29,30]. In this divergence scale, it is almost impossible to identify positively selected and fast evolving genes as regular MS studies do due to a large number or even saturation of substitutions. We, therefore, after obtaining good gene annotation and constructing reliable phylogeny among the lobe-finned lungfish, a comprehensive selection of ray-finned fishes and tetrapods using cartilaginous fishes as the outgroup, counted on identification of lineage-specific conserved elements and newly evolved genes. Together with organ (e.g., lung) gene expression profiles and some experimental validations, we were able to trace the details of genomic evolution along different bony fish lineages, including the one leading to the tetrapods. As pointed out by the Swedish paleontologist, Dr. Per Erik Alhberg [31], previously this kind of research can only be done using fossil record, "but that only tells us about changes in the skeleton. The equally important changes that must have occurred in the soft anatomy and physiology would appear to be lost, undocumented, in the mists of time." Thereby, this vertebrate landing study provides a good example to show that the evolutionary genomics approach crossing large evolutionary scale can bridge the gap between microevolutionary and macroevolutionary studies.

Besides curiosity on pure natural life mysteries, medical applications of extraordinary ability/traits of other organisms are also of great interest to biologists as well as publics. Our recent evolutionary genomics investigations on special ruminant traits provided a promising example in this direction (Figure 1B; Figure 2). Giraffes have 2-fold higher blood pressure than other mammals due to their high stature and thus necessity of pumping blood to the head. After evolutionary genomics analysis against other ruminants and mammals, we identified the *IGFRL1* gene to be the most significant positively selected gene in giraffes, and more interestingly after we replaced the mouse *IGFRL1* gene with the giraffe homolog, the mice not only became avoidant of high blood pressure in the regular hypertension modeling procedure but also acquired higher bone density in adults (Figure 1B) [32]. This case vividly shows that the evolutionary genomics approach proposed in this paper can have great power to reveal molecular mechanism underpinning important traits and enlighten future biomedical research in two major chronic diseases, namely, hypertension and osteoporosis (Figure 1B). In ruminants, the antler of deer is the only mammalian organ that is capable of full regeneration [33], and deer were found to have much lower cancer rate among mammals [34,35]. The reindeer (*Rangifer tarandus*) is an Arctic and subarctic species, in which both females and males can grow big-sized bony antlers in the environment without regular daily sunlight rhythm. Based on well-resolved phylogeny of the ruminant suborder [36], it was revealed that the deer antler shared the same origin as other various ruminants' headgears, but besides sharing basic headgear genes the rapid regeneration growth of antlers also recruited many oncogenic genes, which made its gene expression profile being closest to that of osteosarcoma [37]. Surprisingly, many tumor suppression genes of deer, especially those in the *TP53* pathway, have experienced strong selection, which was proposed to account for the low cancer rate in cervids [38]. As highlighted by Ker and Yang (2019), the "studies of deer antlers offer attractive approaches for tissue engineering and regenerative medicine" [38]. We are now trying to identify if there are special stem cells in the regenerating tissue of antlers, and if there are, it would have great value in regenerative medicine.

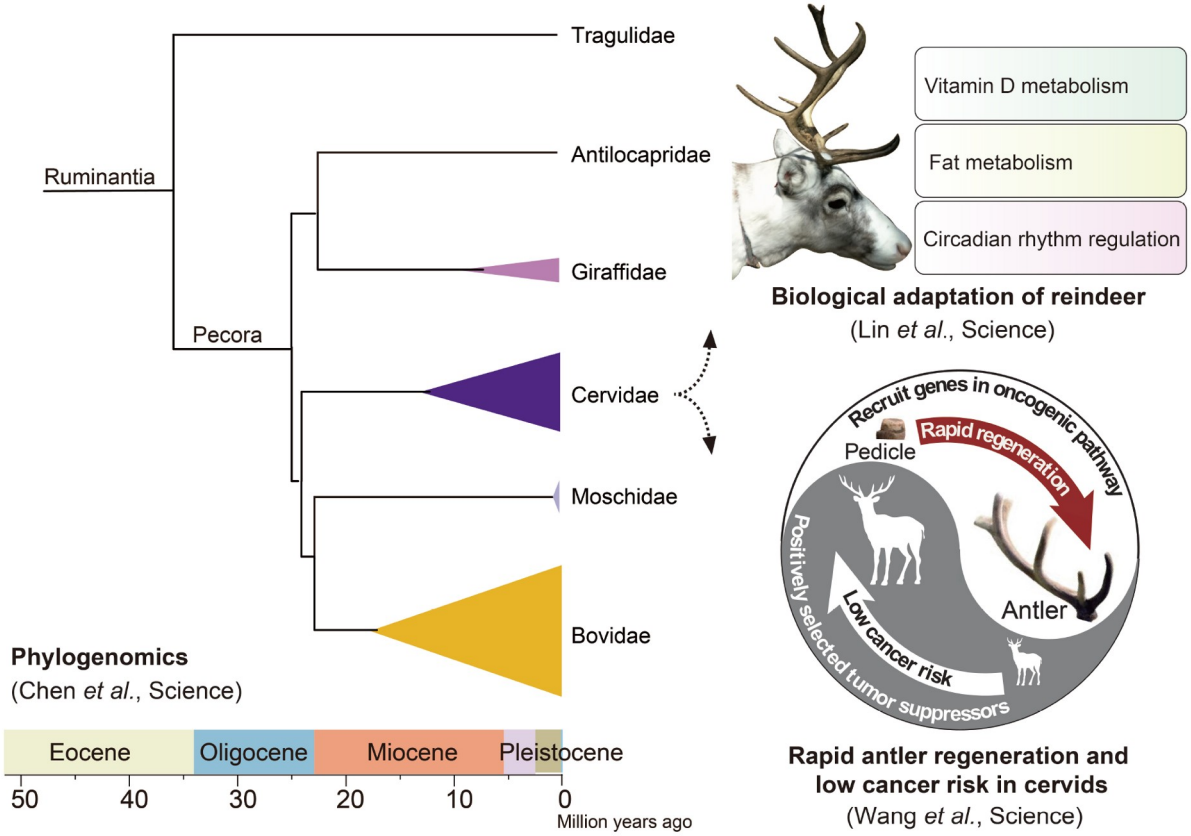

Figure 2 Evolutionary genomics studies on the ruminant suborder shed light on such biomedical research as regenerative medicine, anti-cancer, sleep disorder and osteoporosis.

In reindeer, many circadian genes have species-specific mutations, and experiments showed that one mutation in the key upstream circadian gene *PER2* is sufficient to result in the loss of circadian rhythm [39]. These genomic alternations may be important components in facilitating the adaptation of reindeer to the Arctic arrhythmic light condition. The circadian arrhythmicity in reindeer is correlated to its fragmented sleep while reindeer can energetically migrate thousands of miles in a year, providing a unique research direction on human sleep disorder. It was also found that in reindeer two key vitamin D metabolism enzymes have much higher activity than other mammals, which enable reindeer to procure the high levels of active vitamin D needed to sustain their metabolism and calcium absorption in the Arctic low sunlight environment [39]. This result has enlightening implication in the biomedical research on rickets (osteomalacia) and osteoporosis, a disease widely occurring in old people, especially women.

As for lifespan, through genome comparisons among mammals, scientists have revealed that the long lifespan and low cancer mortality of elephants may result from much more *TP53* genes than other mammals [40]. Evolutionary genomics analyses across mammals and vertebrates also suggest that changes in genes related to DNA repair, cancer and aging may be the underlying molecular mechanisms responsible for the longest-living mammal—bowhead whale [41] and the longest-living vertebrate—giant tortoise [42]. All these recent evolutionary genomics studies with biomedical implications excitingly indicate that the new evolutionary genomics approach has great potential or power to reveal many fantastic life mysteries and

enlighten a novel biomedical research direction.

With the advance of many genome projects including the Earth Biogenome Project (EBP) [43], more and more organisms' genomes will be deciphered, and thus we are facing unprecedented opportunities to uncover miraculous life mysteries evolving on this planet using evolutionary genomics approach. Novel computational methods including such artificial intelligence methods as machine learning are expected to be developed to compare enormous amount of omics data and help identify key genomic elements responsible for the trait innovations. Large-scale and high-throughput experimental validation of these elements will further lay foundation for possible clinic or bionic applications of those superior abilities being learned from other organisms. In the scientific aspect, this will facilitate the long-desired unification of biology [9] by integrating genetics, development and evolution, namely, tightly woven Grand Synthesis of biology, which will make biological traits traceable, predictable and artificially regulatable as we proposed in the eGPS proposal [23].

Under this circumstance, it is urgent to speed up the legislation about ethics of biological research to avoid the appearance of such crazy biologist as Frankenstein, who is a creator of monsters described in an 1818 novel written by the English writer Mary Shelley. It seems that legislation process often seriously lags behind rapid advances of science and technology nowadays. In the fields of artificial intelligence and biotechnology, prompt legislation would not only preventively rule out the possible appearance of Frankenstein, but also be able to promote their positive and proper development.

Acknowledgements

I would like to thank Dr. Qiang Qiu for help in preparing Figure 1, Dr. Lei Chen for help in preparing Figure 2, and Dr. Chunyan Chen for help in editing the manuscript.

Funding

This work was supported by the Talents Team Construction Fund of Northwestern Polytechnical University and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB13000000).

Conflict of interest

The author declares no conflict of interest.

References

1 Darwin C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray, 1859. 489–450.

2 Burnett S. Perceptual worlds and sensory ecology. *Nat Educa Knowledge* 2011; 3: 75.

3 Pris AD, Utturkar Y, Surman C, *et al.* Towards high-speed imaging of infrared photons with bio-inspired nanoarchitectures. *Nat Photon* 2012; **6**: 195–200.

4 Ma Y, Bao J, Zhang Y, *et al.* Mammalian near-infrared image vision through injectable and self-powered retinal nanoantennae. *Cell* 2019; **177**: 243–255.e15.

5 Nelidova D, Morikawa RK, Cowan CS, *et al.* Restoring light sensitivity using tunable near-infrared sensors. *Science* 2020; **368**: 1108–1113.

6 Xu Q, Wan Y, Hu TS, *et al.* Robust self-cleaning and micromanipulation capabilities of gecko spatulae and their bio-mimics. *Nat Commun* 2015; **6**: 8949.

7 Gilbert SF. The Immortal Life Cycle of Turritopsis. In: *Developmental Biology* (9th ed.). London: Oxford University Press, 2010.

8 Schiffman JD, Breen M. Comparative oncology: what dogs and other species can teach us about humans with cancer. *Phil Trans R Soc B* 2015; **370**: 20140231.

9 Smocovitis VB. Unifying biology: The evolutionary synthesis and evolutionary biology. *J Hist Biol* 1992; 25: 1–65.

10 Lamarck JB. *Zoological Philosophy: An Exposition with Regard to the Natural History of Animals*. Chicago: University of Chicago Press, 1984.

11 Dobzhansky T. *Genetics and the Origin of Species*. New York: Columbia University Press, 1937.

12 Mayr E. *Systematics and the Origin of Species from the Viewpoint of a Zoologist*. New York: Columbia University Press, 1942.

13 Stebbins GJ. *Variation and Evolution in Plants*. New York: Columbia University Press, 1946.

14 Kimura M. *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press, 1983.

15 King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. *Science* 1975; **188**: 107–116.

16 Carroll SB. *Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom*. New York: W. W. Norton & Company, 2006.

17 Müller GB. Evo-devo: Extending the evolutionary synthesis. *Nat Rev Genet* 2007; 8: 943–949.

18 Loewe L. A framework for evolutionary systems biology. *BMC Syst Biol* 2009; **3**: 27.

19 Soyer OS, O'Malley MA. Evolutionary systems biology: What it is and why it matters. *BioEssays* 2013; **35**: 696–705.

20 Laland K, Uller T, Feldman M, *et al.* Does evolutionary theory need a rethink? *Nature* 2014; **514**: 161–164.

21 Müller GB. Why an extended evolutionary synthesis is necessary. *Interface Focus* 2017; 7: 20170015.

22 Jablonka E, Lamb M. *Evolution in Four Dimensions: Genetic*, *Epigenetic*, *Behavioral*, *and Symbolic Variation in the History of Life*. Boston: MIT Press, 2014.

23 Chen L, Qiu Q, Pan XY, *et al.* Evolutionary genotype-phenotype systems biology and study on the ruminant evolution. *Sci Sin-Vitae* 2019; **49**: 509–518.

24 Jarvis ED, Mirarab S, Aberer AJ, *et al.* Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 2014; **346**: 1320–1331.

25 Zhang G, Li C, Li Q, *et al.* Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 2014; **346**: 1311–1320.

26 Meredith RW, Zhang G, Gilbert MTP, *et al.* Evidence for a single loss of mineralized teeth in the common avian ancestor. *Science* 2014; **346**: 1254390.

27 Pfenning AR, Hara E, Whitney O, *et al.* Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* 2014; **346**: 1256846.

28 Whitney O, Pfenning AR, Howard JT, *et al.* Core and region-enriched networks of behaviorally regulated genes and the singing genome. *Science* 2014; **346**: 1334.

29 Wang K, Wang J, Zhu C, *et al.* African lungfish genome sheds light on the vertebrate water-to-land transition. *Cell* 2021; **184**: 1362–1376.e18.

30 Bi X, Wang K, Yang L, *et al.* Tracing the genetic footprints of vertebrate landing in non-teleost ray-finned fishes. *Cell* 2021; **184**: 1377–1391.e14.

31 Ahlberg PE. A comparative genomic framework for the fish-tetrapod transition. *Sci China Life Sci* 2021; **64**: 664–666.

32 Liu C, Gao J, Cui X, *et al.* A towering genome: Experimentally validated adaptations to high blood pressure and extreme stature in the giraffe. *Sci Adv* 2021; **7**: eabe9459.

33 Li C, Yang F, Sheppard A. Adult stem cells and mammalian epimorphic regeneration-insights from studying annual renewal of deer antlers. *CSCR* 2009; **4**: 237–251.

34 Lombard LS, Witte EJ. Frequency and types of tumors in mammals and birds of the Philadelphia Zoological Garden. *Cancer Res* 1959; 19: 127–141.

35 Griner LA. *Pathology of Zoo Animals: A Review of Necropsies Conducted Over a Fourteen-year Period at the San*

Diego Zoo and San Diego Wild Animal Park. San Diego: Zoological Society of San Diego, 1983.

36 Chen L, Qiu Q, Jiang Y, *et al.* Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. *Science* 2019; **364**: eaav6202.

37 Wang Y, Zhang C, Wang N, *et al.* Genetic basis of ruminant headgear and rapid antler regeneration. *Science* 2019; **364**: eaav6335.

38 Ker DFE, Yang YP. Ruminants: Evolutionary past and future impact. *Science* 2019; **364**: 1130–1131.

39 Lin Z, Chen L, Chen X, *et al.* Biological adaptations in the Arctic cervid, the reindeer (*Rangifer tarandus*). *Science* 2019; **364**: aav6312.

40 Abegglen LM, Caulin AF, Chan A, *et al.* Potential mechanisms for cancer resistance in elephants and comparative cellular response to DNA damage in humans. *JAMA* 2015; **314**: 1850.

41 Keane M, Semeiks J, Webb AE, *et al.* Insights into the evolution of longevity from the bowhead whale genome. *Cell Rep* 2015; **10**: 112–122.

42 Quesada V, Freitas-Rodríguez S, Miller J, *et al.* Giant tortoise genomes provide insights into longevity and age-related disease. *Nat Ecol Evol* 2019; **3**: 87–95.

43 Lewin HA, Richards S, Lieberman Aiden E, *et al.* The Earth BioGenome Project 2020: Starting the clock. *Proc Natl Acad Sci USA* 2022; **119**: e2115635118.