Perspectives on studying molecular adaptations of amphibians in the genomic era

Yan-Bo Sun[1,2,*], Yi Zhang[2], Kai Wang[2,3]

[1] *Laboratory of Ecology and Evolutionary Biology, Yunnan University, Kunming, Yunnan 650091, China*

[2] *State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China*

[3] *Sam Noble Oklahoma Museum of Natural History and Department of Biology, University of Oklahoma, Norman, Oklahoma 73072, USA*

ABSTRACT

Understanding the genetic mechanisms underlying particular adaptations/phenotypes of organisms is one of the core issues of evolutionary biology. The use of genomic data has greatly advanced our understandings on this issue, as well as other aspects of evolutionary biology, including molecular adaptation, speciation, and even conservation of endangered species. Despite the well-recognized advantages, usages of genomic data are still limited to non-mammal vertebrate groups, partly due to the difficulties in assembling large or highly heterozygous genomes. Although this is particularly the case for amphibians, nonetheless, several comparative and population genomic analyses have shed lights into the speciation and adaptation processes of amphibians in a complex landscape, giving a promising hope for a wider application of genomics in the previously believed challenging groups of organisms. At the same time, these pioneer studies also allow us to realize numerous challenges in studying the molecular adaptations and/or phenotypic evolutionary mechanisms of amphibians. In this review, we first summarize the recent progresses in the study of adaptive evolution of amphibians based on genomic data, and then we give perspectives regarding how to effectively identify key pathways underlying the evolution of complex traits in the genomic era, as well as directions for future research.

INTRODUCTION

Evolution can be seen as the accumulations of species' adaptations to external environments (Dobzhansky & Gould, 1982), thus understanding the mechanisms underlying the organisms' adaptations, especially at the molecular level, has been one of the core issues of evolutionary biology. Moreover, such understanding can also provide an efficient framework to reveal the relationship between genotype and phenotype (i.e., forward genetics; Figure 1). Through years of field and laboratory observations and measurements, a large number of cases of phenotypic adaptations, both morphological and physiological, have been identified in vertebrates, such as the limb evolution and their corresponding adaptations to specific locomotion types in bats and whales (Liang et al., 2013), the evolution of antimicrobial peptides and adaptations to amphibious skin structures in frogs (Rollins-Smith, 2009), and
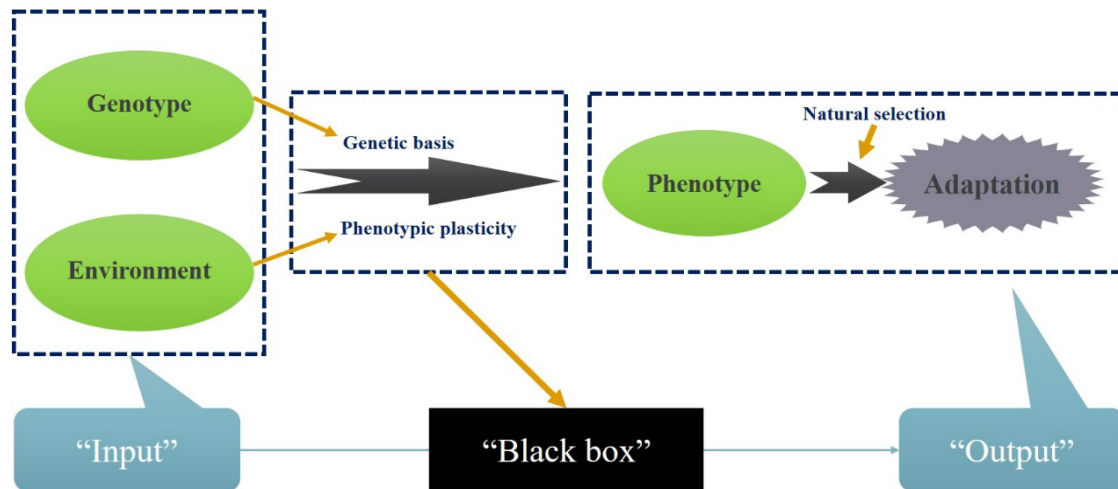
Figure 1 The logical framework of adaptive evolution study

Genotype plus environment result into particular phenotype, which then experienced diverse roles of natural selection and finally formed adaptation. During this process, the genotype forms the genetic basis of phenotype, while the environment promotes the formation of phenotypic plasticity. Under this framework, given specific genotype and environment ("input"), one can get a certain phenotype ("output"). Between input and output, the genetic mechanisms underlying phenotype formation and the source of phenotypic plasticity constitute a "black box", during which the study of adaptive evolution becomes one efficient tool to open the box.

the most classic evolution of the beak sizes/shapes of Darwin's finches and the adaptation to specific ecological niches (Lamichhaney et al., 2015). These case studies provide rich materials for the study of adaptive evolution and the underlying molecular mechanisms.

With the recent technical advances, large quantities of DNA or RNA can be sequenced in a much more efficient way. In addition to the *de novo* whole-genome sequencing, the recent advances in sequencing technologies (i.e., resequencing, RNA-seq, and restriction site-associated DNA sequencing) and computational power further enable the genomic era. It has been proposed that the new genomic data can provide great insights for a diverse set of questions in evolutionary biology, such as questions on phylogenetic relationships and the tree of life, genome-size evolution, historical demography and current population structure, adaptive potential, hybridization and speciation, and the genetic bases of phenotypic traits (Allendorf et al., 2010; Brandies et al., 2019; Liedtke et al., 2018; Supple & Shapiro, 2018). Under this view, more and more vertebrate genomes, transcriptomes and other omics data are accumulated and now available, even for some non-model species (https://www.ncbi.nlm.nih.gov/genome). However, the availability of genomes is still heavily taxon-biased. Until now, majority of the available vertebrate genomes are from mammals and birds (Allendorf, 2017; Genome 10K Community of Scientists, 2009; Ostrander et al., 2019; Supple & Shapiro, 2018; Zhang, 2015), and for other diverse vertebrate groups, only very few genomes were available comparatively. For amphibians specifically, of the 8 043 currently recognized amphibian species (https://amphibiaweb.org/index.html, 2019; Che & Wang, 2016), no more than 20 of them have published or released genomes

available (Table 1). Since amphibians possess many unique characteristics for the study of genome evolution and molecular adaptations (see below), such lack of genomic information really limits our understandings of amphibian evolution, as well as conservation (Funk et al., 2018; Shaffer et al., 2015).

One major reason leading to such "genome deficient" in amphibians may come from the methodological challenges for assembling very large, repetitive genomes (Elliott & Gregory, 2015). It has been reported that the estimated average genome sizes of Anurans (frogs), Gymnophionas (caecilians), and Caudatas (salamanders) were 4.1, 5.6, and 32 gigabases (Gb), respectively (Liedtke et al., 2018). Such large genomes of amphibians post major challenges to both the sequencing and the assembling processes of the genomic data. However, with the fast development of sequencing technology, particularly the "third" (i.e., PacBio's single molecule real-time system) and even "fourth" generation of sequencing techniques (i.e., Oxford Nanopore PromethION system) (Deamer et al., 2016; Glenn, 2011), as well as the efficient assembly methods (Ruan & Li, 2020), such difficulties could now be overcome to a large extent. As results, more and more high-quality genome assemblies of amphibians are emerging in the recent years (Li et al., 2019b; Nowoshilow et al., 2018; Smith et al., 2019).

For the approaches used to determine the genetic mechanisms underlying trait evolutions, two main approaches were used in both comparative (interspecific) and population (intraspecific) genomic approaches utilize commonly used methods (i.e., d_N/d_S for comparative and F_{ST} for population genomics) through genome-wide scans for loci under positive selection in particular lineages or populations (Lee & Coop,

Table 1 List of amphibian species with available reference genome

Order	Family	Species	Common name	Genome size (Gb)	Scaffold N50 (Mb)	Contig N50 (Kb)	Assembly level	Sequencing technology	References	Link
Gymnophiona	Siphonopidae	*Microcaecilia unicolor*	Common caecilians	4.68	376.14	3 660	Chromosome	PacBio	N/A	https://vgp.github.io/genomeark/Microcaecilia_unicolor/
Gymnophiona	Dermophiidae	*Geotrypetes seraphini*	Gaboon caecilian	3.78	272.61	20 660	Chromosome	PacBio	N/A	https://vgp.github.io/genomeark/Geotrypetes_seraphini/
Gymnophiona	Rhinatrematidae	*Rhinatrema bivittatum*	Two-lined caecilian	5.32	486.88	0.350	Chromosome	PacBio	N/A	https://www.ncbi.nlm.nih.gov/assembly/GCF_901001135.1/
Caudata	Ambystomatidae	*Ambystoma mexicanum*	Axolotl	32.4	1410	210	Chromosome	PacBio	Nowoshilow et al., 2018	https://www.ncbi.nlm.nih.gov/assembly/GCA_002915635.2/
Caudata	Salamandridae	*Pleurodeles waltl*	Iberian ribbed newt	19.38	N/A	N/A	Chromosome	Illumina	Elewa et al., 2017	https://www.ncbi.nlm.nih.gov/bioproject/353981
Anura	Hylidae	*Dendropsophus ebraccatus*	Hourglass treefrog	2.24	60.91	8860	Chromosome	PacBio	N/A	https://vgp.github.io/genomeark/Dendropsophus_ebraccatus
Anura	Megophryidae	*Leptobrachium leishanense*	Leishan spiny toad	3.55	394.69	1 900	Chromosome	PacBio RS II	Li et al., 2019a	https://www.ncbi.nlm.nih.gov/genome/46619?genome_assembly_id=743320
Anura	Megophryidae	*Vibrissaphora ailaonica*	Ailao spiny toad	3.53	412.42	820	Chromosome	PacBio	Li et al., 2019b	http://gigadb.org/dataset/100624
Anura	Ranidae	*Lithobates catesbeianus*	American bullfrog	5.8	0.05	5	Scaffold	Illumina	Hammond et al., 2017	https://www.ncbi.nlm.nih.gov/genome/230312genome_assembly_id=353399
Anura	Ranidae	*Rana temporaria*	Common frog	4.18	0.05	2.889	Scaffold	Illumina HiSeq	N/A	https://www.ncbi.nlm.nih.gov/assembly/GCA_009802015.1
Anura	Pyxicephalidae	*Pyxicephalus adspersus*	African bullfrog	1.56	157.52	30	Chromosome	Illumina	Denton et al., 2018a	https://www.ncbi.nlm.nih.gov/assembly/GCA_004786255.1/
Anura	Bufonidae	*Rhinella marina*	Marine toad	2.55	N/A	160	Contig	PacBio RS II	Edwards et al., 2018	https://www.ncbi.nlm.nih.gov/assembly/GCA_900303285.1/
Anura	Scaphiopodidae	*Scaphiopus couchii*	Couch's spadefoot toad	0.48	N/A	0.362	Scaffold	Illumina	Seidl et al., 2019	https://www.ncbi.nlm.nih.gov/genome/851442genome_assembly_id=723961
Anura	Scaphiopodidae	*Scaphiopus holbrookii*	Eastern spadefoot toad	0.71	N/A	0.514	Scaffold	Illumina	Seidl et al., 2019	https://www.ncbi.nlm.nih.gov/genome/690972genome_assembly_id=723960
Anura	Scaphiopodidae	*Spea bombifrons*	Plains spadefoot toad	0.77	N/A	0.522	Scaffold	Illumina	Seidl et al., 2019	https://www.ncbi.nlm.nih.gov/assembly/GCA_009364475.1/
Anura	Scaphiopodidae	*Spea multiplicata*	Mexican spadefoot toad	1.07	0.07	30	Scaffold	PacBio	Seidl et al., 2019	https://www.ncbi.nlm.nih.gov/assembly/GCA_009364415.1/
Anura	Dicroglossinae	*Nanorana parkeri*	Tibetan frog	2.05	1.06	30	Scaffold	Illumina	Sun et al., 2015	https://www.ncbi.nlm.nih.gov/assembly/GCF_000935625.1/
Anura	Pipidae	*Xenopus laevis*	African clawed frog	2.72	136.57	20	Chromosome	Illumina	Session et al., 2016	https://www.ncbi.nlm.nih.gov/assembly/GCF_001663975.1/
Anura	Pipidae	*Xenopus tropicalis*	Tropical clawed frog	1.45	153.96	14 630	Chromosome	PacBio; Illumina HiSeq	Session et al., 2016	https://www.ncbi.nlm.nih.gov/assembly/GCF_000004195.4/
Anura	Dendrobatidae	*Oophaga pumilio*	Strawberry poison frog	5.5	0.072	0.385	Scaffold	Illumina HiSeq	Rogers et al., 2018	https://www.ncbi.nlm.nih.gov/genome/86474

N/A: Not available.

2019; Zhang et al., 2014). Generally, comparative genomics can reveal the evolutionary patterns of genes within a large time scale and determine whether the genes experienced rapid or slow evolution, while the population genomics can identify genes associated with local adaptation of one or more populations in a relatively short time scale. One main logic behind the two approaches is to search for outlier genes that show signals of strong selection and are significantly separated from the background genes. At present, there have been lots of example studies having applied such analytical logic. For instance, the recent analyses of dozens of ruminant genomes makes it possible to decipher the genetic underpinnings of multiple phenotypes of this taxa, like the evolution of headgear and multichambered stomach, and thus makes this taxa a good model for further genomic analyses, like studying adaptive evolutionary mechanisms of them (Wang et al., 2019).

Under these analytical frameworks, the genomic analyses that have been used commonly in mammals and other well-studied groups are gradually applied into amphibians, leading to important findings on the macro- and adaptive evolution of amphibians at the molecular level (Nowoshilow et al., 2018; Sun et al., 2015, 2018; Wang et al., 2018; Yang et al., 2012). However, because of the specific characteristics of amphibian genomes (i.e., high repeatability and incomplete annotations of genomic elements), simply copying the analytical methods of the past may not be sufficient to reveal the evolutionary mechanisms of adaptive evolution of these unique organisms. With the accumulation of omics data, evolutionary herpetologists need to beware of analytical methods used and be able to interpret the results from genomic dataset.

There have been several reviews that focus on the applications of omics data to better understand amphibian conservation, ecology, and evolution (Funk et al., 2018; Shaffer et al., 2015; Storfer et al., 2009), yet none of them focuses on the evolutionary patterns and mechanisms of amphibians' adaptive evolution, or how to apply the accumulated omics data to study amphibians effectively. In this review, we first give a brief overview of recent progresses of studies on the adaptive evolution of amphibians, and then we focus our discussions by giving perspectives on the future directions for studies on adaptive evolution of amphibians, including the potential contributions of the repetitive elements on the genome evolution of amphibians.

WHAT WE HAVE LEARNT FROM THE AVAILABLE AMPHIBIAN GENOMES?

As shown in Table 1, to date, a total of 20 amphibian genome sequences have been released (Denton et al., 2018; Edwards et al., 2018; Elewa et al., 2017; Hammond et al., 2017; Li et al., 2019a, 2019b; Nowoshilow et al., 2018; Rogers et al., 2018; Seidl et al., 2019; Session et al., 2016; Sun et al., 2015). Most of these genomes are from Anurans (*n*=15), and there are only two and three complete genomes available for Caudata and Gymnophiona, respectively (Table 1).

Considering the diversity of amphibians, the current state of genomic research on amphibians is much behind other taxonomic groups. The majority of these genomes were used to explore the associations between phenotypes and genotypes and the possible values of these phenotypes to the evolutionary adaptations. For instance, the recent genome of *Leptobrachium leishanense* was used to explore the genetic bases of sexually dimorphic traits (Li et al., 2019a), and the large genome of Axolotl (*Ambystoma mexicanum*) was used to explore the key regulators of tissue regeneration (Nowoshilow et al., 2018). Phenotypic evolution itself is a diverse research field, which covers important topics on the compositions of genetic architecture, phenotypic variance and heritability, phenotypic correlation, sources and costs of phenotypic plasticity, as well as phenotypic adaptation. On the other hand, by providing both the prophase theoretical fundaments and downstream functional validations, results from phenotypic evolution (i.e., the evolutionary and developmental mechanisms of the phenotypes) can provide crucial insights for the study of molecular adaptation. Conversely, the analytical methods developed in studying the molecular adaptation can also be suitable for determining the genetic architecture of the interested phenotypes. Here, we give a brief review of the latest literatures on identifying the molecular adaptations of the amphibians.

Amphibians evolve slowly, but still have significant molecular adaptations

With the assembly of the first "modern" frog genome (that of *Nanorana parkeri*), Sun et al.. (2015) gave the first glimpse of the evolutionary rate of frogs. Through the whole-genome comparisons between *N. parkeri* and *Xenopus tropicalis*, the study showed that the ectothermic vertebrates (i.e., amphibians, lizards, and fishes) have significantly slower evolutionary rates than endothermic vertebrates, which are indicated by less structural variation and slower base substitution rates (Sun et al., 2015). Such conclusion about the slower evolutionary rate was further evidenced when the third frog genome (of *Lithobates catesbeianus*) became available (Hammond et al., 2017).

It is known that comparative genomics allow a better understanding of the evolutionary patterns of genes. However, it is generally difficult to perform genomic/gene comparisons of different taxa if the two species differentiate for a long time, as the long-term accumulation of base substitution posts challenges due to saturation. Furthermore, given the positive selection often operates episodically in a short timescale, it can be masked by the long-term effects of purifying selection (Zhang et al., 2005). Hence, to identify the candidate loci underlying particular traits, comparative genomics is generally conducted among closely related and recently diversified species (Wang et al., 2019). For amphibians, their slower evolutionary rate at the molecular level provides two potential advantages in studying evolutionary patterns of genes through a comparative approach: (1) it could be easier to identify loci that hold significant differentiations between pair of taxa or

populations with shallow divergence, in the case of much similar genomic background; and (2) it could also allow genomic comparisons between distantly related, deeply diverged species or populations. These two advantages make amphibians an excellent group to study the genetic bases underlying their rich phenotype adaptations. In addition to the rich phenotypic diversity (i.e., body size, coloration, respiratory rate, and toxicity), studying the genotype-phenotype associations in amphibians would become a research hotspot in the future. For example, the frogs inhabiting high-elevation environments evolved numerous phenotypic adaptations, such as the higher secretion of antioxidant peptides (Yang et al., 2016) and higher numbers of epidermal capillaries and granular glands (Yang et al., 2019), which providing excellent models to investigate their underlying genetic mechanisms for high-elevation adaptation.

To overcome the limitation in the availability of genomes for amphibians, comparative transcriptomics methods that compare genome-wide transcribed RNA sequences are an alternative and useful option (Wang et al., 2009). The comparative transcriptomic approach can identify genes under selections through the existing methods, without a fully assembled reference genome. Such approach has been already adopted in the recent studies of adaptations in amphibians, particularly on the genetic basis of phenotypic adaptations in high-elevation frogs (Sun et al., 2018; Yang et al., 2012). Yang et al. (2012) found the *Rana* species at high-elevation experienced rapid evolution at genes response to hypoxia and oxidative stress. Sun et al., (2018) further compared four species of frogs (three plateau- and a lowland-species of *Nanorana*) and lizards (*Phrynocephalus*) across an elevation gradient to examine the gradual accumulation of high-elevation adaptations. By identifying signals of positive selections along the phylogeny, Sun et al., (2018) found that numerous molecular adaptations to high elevations, especially the DNA-repair and energy-metabolism pathways, appear to arise gradually and evolve continuously as the elevation increases. Such results suggest crucial roles of the two functional pathways during the adaptations to high elevations in frogs. Furthermore, Sun et al., (2018) also provide a unique and efficient comparative framework for studying the adaptive evolution by using multiple species in different groups that distributed across the same environmental gradients.

Population genomics help identifying candidate loci underlying local adaptations

The slower evolutionary rates of amphibians can be more beneficial to population genomics study, which compares different populations of a same species with more similar genomic backgrounds. Given the relatively low dispersal ability and the resulting limited gene flows among amphibian species & populations (Ward et al., 1992), local adaptations should be a common phenomenon in amphibians (Funk et al., 2018), especially for amphibian species spanning dramatic environmental gradients (Funk et al., 2016; Wang et al., 2018). Hence, conducting population genomic studies on

amphibians have its unique advantages over other taxonomic groups.

Multiple analytical methods have been developed to detect the candidate loci that are associated with the local adaptations of interests, and these methods can be classified into three groups based on the sequencing method and the data type (Nadeau & Jiggins, 2010). The first group identifies loci with divergence higher than those observed at neutral loci, and it uses the data from whole-genome resequencing (Beaumont & Balding, 2004). However, because the quality of analyses for this group of methods largely depends on the availability of reference genome, hence, up to now, this kind of study is still in its infancy in amphibians due to the lack of reference genomes. Recently, Wang et al., (2018) conducted the first-ever whole-genome resequencing study in amphibians. This study took advantage of the published reference genome of the *N. parkeri* (Sun et al., 2015) and sequenced 63 new individuals to infer the historical demography, speciation, hybridization, and potential genomic bases of adaptation to high elevations of the higher-elevation populations (Wang et al., 2018). Their results showed that natural selection plays important roles in driving and maintaining the continuing divergence within *N. parkeri*, and the results identified several candidate genes (e.g., *CAT* (Catalase), an antioxidant protein coding gene) that show high-divergence in both genetic sequences and expression levels of genes between the high- and low-elevation populations (Wang et al., 2018), which further evidence the evolution of molecular adaptations in amphibians.

The second group of methods refers to the restriction-site associated DNA sequencing (RADseq), which provides a reduced representation approach to population genomic studies (Andrews et al., 2016). With this method, a greater number of individuals can be genotyped at a smaller number of loci in the genome. More importantly, it can be used independently of the reference genome, although a reference genome will still be required to identify specific gene of interests. Generally, the data generated from RADseq are enough to answer questions regarding population structure and demographic history (Andrews et al., 2016). Recently, Guo et al., (2016) used genome-wide scans of SNPs from RADseq data to identify loci under strong divergent selection between populations of *Bufo andrewsi* at low- and high-elevations, and they speculated some of these SNPs are associated with differences in elevation/temperature and are involved in adaption to high elevations (Guo et al., 2016). However, without a reference genome, it is difficult to determine which specific genes do the identified SNPs come from.

The third group of methods to study adaptive divergence associated with local adaptation is to perform population transcriptomics using RNAseq (Wang et al., 2009). Comparing to the previous two groups of methods, although the transcriptomic methods imposes more challenges especially in sample collection and preservation, they can provide quantitative expression levels of individual genes, in addition

to the typical information on the DNA sequences, which has unique advantages in answering questions on adaptive evolution at the molecular level. Despite these advantages, to date no published work has taken the population transcriptomic approach in studying amphibian adaptations. Recently, our laboratory has explored the transcriptomic dataset on amphibians, and our preliminary results confirm the promising future of this group of method in studying adaptive evolution in amphibians. We constructed the expression matrix of more than 10 000 protein-coding genes, using the RNA transcriptomics samples from five populations of *N. parkeri* through an elevation gradient (from 2 800 m to 4 800 m a.s.l.) from the Tibetan Plateau. Based on this expression matrix, we identified numerous of candidate genes whose expression shifts are significantly correlated with the elevations, which may provide another way to study the molecular adaptation of amphibians (unpublished data).

Molecular convergence is present between amphibians and distantly related species

Convergent evolution is defined as the process whereby two distantly related species evolve similar traits independently as the result of similar selective forces, rather than shared due to common ancestry (Böcher, 1977). Although such phenomena are mostly known at macro-level (i.e., phenotypic convergence), like the flight ability of mammals and avian, and the ultrasonic communication in amphibians and malmmals (Shen et al., 2008), the macro-level convergences are the results of micro-level convergences (i.e., molecular convergence), which are often overlooked. For instance, the high-frequency acoustic sensitivity and selectivity of bat and whale echolocation rely on a common molecular design of *prestin* gene, where occurred some common amino-acid substitutions between the two echolocation species (Li et al., 2010).

Although most studies on molecular convergence focused on more closely related species/lineages (Chikina et al., 2016; Foote et al., 2015; Li et al., 2008; Thomas & Hahn, 2015), we can still draw conclusions about the presence of molecular convergence between distantly related species based on comparisons of these available studies. One classical example is the evolution of *EPAS1* gene in multiple independent plateau lineages. The *EPAS1* gene encodes a transcription factor involved in the induction of hypoxia regulating genes (Tian et al., 1998). Rapid evolutions and positive selections of *EPAS1* are documented in varies plateau organisms, from short-term evolution in the human Tibetan population (Beall et al., 2010), to the long-term evolution of Tibetan Hot-Spring Snakes of the genus *Thermophis* (Li et al., 2018). Other examples include different DNA-repair genes. Some of these genes have been found undergoing positive selection in plateau frogs (Yang et al., 2012) and plateau mammals (Ge et al., 2013). Furthermore, 32 homologous genes were found under positive selections in both frogs and lizards in Tibet, many of which are responsible for DNA-repair and energy-metabolism (Sun et al., 2018).

Nonetheless, the molecular convergence mainly occurred at pathway/network level rather than at the individual gene or site level (Hao et al., 2019; Sun et al., 2018). Moreover, between different groups, molecular convergence may occur in different evolutionary stages (Sun et al., 2018), which suggest that sampling multiple species across environmental gradients will provide more power for testing molecular adaptations than focusing on individual species at a single locality (Solak et al., 2020). Furthermore, as more convergences occur at the pathway/subnetwork level, it suggests that adaptive traits are commonly controlled by multiple genes, and the traditional methods through detecting outlier genes may not be enough effective in detecting the signals and recovering the full story.

FUTURE PERSPECTIVES ON STUDYING ADAPTIVE EVOLUTION OF AMPHIBIANS

Although we started to get a grasp on the genomic evolution and the molecular mechanisms underlying their adaptive traits in amphibians, there are still numerous limitations that need to be solved before we gain a full understanding. Firstly, amphibians are extremely diverse in terms of their life history, reproductive mode, habitat, genome size, and interaction with other species, for which the currently available genomes are still not enough to fully understanding their diverse evolutionary adaptations and lineage-specific traits, like how some frogs adapt to the marine environment (i.e., *Fejervarya multistriata*) and what the forming mechanisms underlying the flippers of tree frogs. Further, another interesting question refers to what the pattern of natural selection pressures acting on different genomic regions of the amphibians with their diverse life styles. For example, whether the different reproduction modes of amphibians imposed different selection pressures acting on the reproduction associated genes, and then drove the different evolution rates of these genes among related species. Hence, the ability to generate more amphibian genomes is perhaps the most important factor affecting future amphibian genome research. With the rapid development of sequencing technology as well as assembling method, the above limiting factor would be solved to the largest extent in the next years.

Secondly, it is inevitable for us to overcome many technical challenges associated with genomics and bioinformatics. During the process of accumulating genomic data, amphibian scientists should pay more attentions on selecting the best suitable species and/or populations and learning more effective analytical methods to address the classic and/or newly raised evolutionary questions, especially those peculiar to amphibians. Given the characteristics of amphibian genomes, like the poor annotations of genomic elements, high repeatability with abundant transposable elements, and much less omics data in terms of expression or methylation, more novel bioinformatic methods and omics data should be developed and generated in the future. For example, the traditional natural selection detection methods are generally performed at the gene level, whose logic is to decide whether

the current selection signal is significantly separated from those of background genes ("outlier" method). However, since complex traits are always controlled by multiple genes, such method may fail to identify the selection signals when the traits are controlled by multiple minor-effect genes rather than by a major-effect gene. For this problem, it becomes necessary to develop some different detection methods, like those based on pathway/network levels.

Here, we provide perspectives on the future studies of amphibians' adaptation and give technical suggestions to determine the genotype-phenotype associations more effectively in the study of amphibians' adaptive evolution.

More closely-related genomes with detailed annotations are necessary

Detecting positive selection at the molecular level is a commonly used method to identify the molecular adaptation. However, as positive selection often operates episodically on a few amino acid sites, it can be difficult to detect under the long-term effects of negative selection (Zhang et al., 2005). Hence, comparing the genes and/or their regulatory sequences of multiple closely-related species will be more efficient in detecting signals of positive selection and also studying evolutionary mechanisms of the adaptive traits. However, because the amphibian genomes are typically large and difficult to assemble, few genomes are available, and the resulting studies are limited greatly for in-group comparisons about adaptive evolution. In the future, with the implementation of various whole-genome sequencing projects, this situation will change. For example, the Genome 10K consortium has provided several criteria for prioritizing amphibian species for whole genome sequencing, which consider the endangered species first and then followed by the genomic "outposts" lineages across the tree of life (Genome 10K Community of Scientists, 2009). According to these criteria, Genome 10K has recommended that each major taxonomic group should have a representative of a very high-quality reference genome (Koepfli et al., 2015), on the basis of which people can perform more detailed comparisons through sequencing the transcriptomes of related species. Specifically, as the cost of genome sequencing drops dramatically, the frog species, which have much smaller genomes than other two orders of amphibians, could become the focus of future genome sequencing projects. It has been estimated that there are about 31% of frog species (n>2 000) having a genome size smaller than that of human (Funk et al., 2018; Liedtke et al., 2018). Considering that the genome size of frogs is relatively conservative within a given clade (Liedtke et al., 2018), sequencing multiple lineages and conducting interspecific comparisons maybe achievable in the near future.

In addition to the availability of comparable genomes, the quality of genome annotation is also crucial for further evolutionary studies. Genome annotation is a key process that identifies the gene properties (i.e., coding vs. non-coding), gene locations, and gene functions. With the development and application of transcript-based sequencing methods, such as

the RNA-seq (Wang et al., 2009) and the *de novo* gene prediction methods (Stanke et al., 2004), the accuracy of predicting gene structures of a given genome has been improved greatly. From the currently available amphibian genomes, the number of protein-coding genes has been estimated to be between 20 000 and 30 000 (Edwards et al., 2018; Hammond et al., 2017; Hellsten et al., 2010; Rogers et al., 2018; Sun et al., 2015). However, unlike the model vertebrate species (i.e., human and mouse), it is still not clear how many non-coding genes are present in amphibian genomes and what biological functions do they play for the development or adaptation. Even for the protein-coding genes, the molecular functions of a larger proportion of them are still unknown in amphibians (i.e., functions of the novel genes). Incomplete genome annotations lead to uncertainty in determining the genetic basis of phenotypic traits. For example, even though studies show that the Tibetan frog (*N. parkeri*) had evolved multiple molecular adaptations to the plateau environments (Sun et al., 2015, 2018; Wang et al., 2018), it is still unclear whether these adaptations benefit from novel genes, and what the relative roles of gene mutations and the gene regulatory/interaction network play in amphibian adaptations. All these questions will be an important research direction in the future, and the answers to them will largely depend on the integrity of genome sequence as well as the level of details in the functional annotations of the genome. This will need years of studies from multiple labs, like the Encode project (The ENCODE Project Consortium, 2012).

Improving the quality of sequence clusters for inter-comparisons

One commonly used method to determine the selection pressure is to estimate the nonsynonymous to synonymous substitution rate ratio (d_N/d_S) of genes under a phylogenetic framework, with the d_N/d_S >, =, and <1 indicating the gene undergone positive selection, neutral evolution, and purifying selection, respectively. As mentioned before, positive selection can be difficult to detect because it often operates episodically on a few amino acid sites, and the signal may be masked by the long-term effect of negative selection (Zhang et al., 2005). This is why the improved branch-site model (Zhang et al., 2005) became the popular method to detect positive selection that affects a small number of sites along specific lineages. However, this model can be easily affected by the quality of multiple sequence alignment (Fletcher & Yang, 2010). Therefore, improving the alignment quality becomes a crucial step before performing tests for positive selection using the branch-site model.

For multiple sequence alignment, it not only contains alignment errors, which may be introduced by aligner and software, but it could also come from primary sequence errors, which may be caused by frame-shifts or erroneous gene annotations. The latter kind of error can have greater impacts on the final alignment quality. For amphibian genomics, their genomes are rich of transposable elements, which can inadvertently insert into both the coding and non-coding

regions, resulting as the primary sequence changes of the host. Such primary sequencing errors are fundamentally different from the alignment errors, which can introduce non-homologous segments in the gene alignments, and such noises are unlikely to be removed by the available trimming methods (Di Franco et al., 2019). Unfortunately, although numerous efforts have been made to reduce alignment errors, only a few recent studies paid attentions to the impacts of primary sequence errors on the alignment construction and detections of positive selection (Di Franco et al., 2019; Whelan et al., 2018). Currently, an effective pipeline for preparing high-quality alignments can be achieved by using a good alignment construction like Prank (Löytynoja & Goldman, 2005), an aligner showing much higher accuracy than others (Fletcher & Yang, 2010), and using an alignment trimming software with one or more filtering methods, like Prequal (Whelan et al., 2018), HmmCleaner (Di Franco et al., 2019), and FasParser (Sun, 2017, 2018). Future studies should give more attentions on the issues of sequencing errors and how to better handle this issue using bioinformatics.

Detecting natural selection on gene subnetworks

Most complex traits are commonly controlled by multiple genes scattered throughout the whole genome (Marouli et al., 2017; Wood et al., 2014), while the traditional ortholog-based methods always treat loci independently and aim to detect outlier genes as candidate genes. Some people even argued that adaptation events occur through the evolution of polygenetic traits rather than via the fixation of single beneficial mutations (Daub et al., 2013; Field et al., 2016). Hence, the caveat of classical genome scans for selection, which mainly focused on identifying the major-effect genes, is that they are inefficient in connecting a list of candidate genes to complex mechanisms of adaptation, and they cannot identify contributions of multiple small-effect genes that contribute to the evolution of traits (Gouy et al., 2017; Vitti et al., 2013). Hence, analyses based on the gene-network, rather than at the ortholog level, can provide a more powerful framework in studying the evolution of adaptive traits and facilitate the interpretation of genome-wide data (Guo et al., 2019; Yu et al., 2017).

Several statistical inference approaches have been proposed to detect selections that act on the polygenic traits (Coop et al., 2010; Hancock et al., 2010; Orr, 1998). On the bases, two feasible solutions have been proposed to identify signals of selection on multiple genes. One is to treat a set of genes as a whole unit (i.e., biological pathways) and test whether the unit undergone any selection (Daub et al., 2013; Foll et al., 2014). The idea of such approach is to assign a score to each gene within a pathway and to then test whether the distribution of scores within the pathway is significantly different from the background (Daub et al., 2013). This approach has been used to identify candidate pathways involved in human response to pathogens (Daub et al., 2013) and adaptations to the high elevations (Foll et al., 2014). However, since this approach only considers the pathways

where all their genes experience a shift in the score distribution, it would be underpowered in identifying more subtle signals, where only a small subset of genes within a large pathway are under directional selection (Gouy et al., 2017). Thus, the second approach is proposed to search for subnetworks of interacting genes within the biological pathways that present unusual features (Gouy et al., 2017). This second approach can discover small-effect genes in complex selective processes, and thus would be an important complement to the classical genome scans. It has been successfully used to identify pathways related to high-elevation adaptations (Gnecchi-Ruscone et al., 2018), as well as response to diseases (Prohaska et al., 2019), showing its applicability in studying amphibian adaptations.

Detecting natural selections on gene expressions across the phylogeny

It has long been accepted that divergence of gene regulation, manifested by changes in gene expression, play a key role in phenotypic evolution (Ferea et al., 1999; Fraser et al., 2010; King & Wilson, 1975), and examining comparative expression levels can help us to identify fundamental changes underlying adaptations to environmental factors (Chen et al., 2019). However, in amphibians, the data generated through RNA-seq technology are commonly used to obtain the protein-coding sequences only (Sun et al., 2018; Yang et al., 2012), and little attention was paid to the gene expression profiles and the roles of expression shift in adaptive evolution.

On major reason for this phenomenon might be the lacks of suitable computational models to detect the pressure of natural selection at the level of gene expression. The level of gene expression can be treated as the traditional quantitative data. Under this view, the quantitative phylogenetic methods that study traditional morphological trait evolution by accounting for nonindependence relationships between species (Felsenstein, 1985; Hansen, 1997; Rohlf, 2001) could be used in analyzing the evolution of gene expressions. However, since both genetic bases and environmental factors can influence gene expression levels (Idaghdour et al., 2009; Pickrell et al., 2010), the changes in expression level may not actually reflect genetic adaptation, and a large proportion of them might be neutral for the host survive (Yang et al., 2017). Thus, due in part to a lack of agreement for how to best model evolution of expression, there was no consensus on a quantitative framework for addressing this issue (Rohlfs & Nielsen, 2015).

This situation has been alleviated in some extent with the proposition of phylogenetic ANOVA, the expression variance and evolution (EVE) model (Rohlfs & Nielsen, 2015). This method treats the gene expression data as a quantitative trait that evolves over a phylogeny and incorporates classic genetic neutrality tests by considering both expressional polymorphism and diversity (Rohlfs & Nielsen, 2015). In the EVE model, the ratio of polymorphism within species to diversity between species is estimated for each gene, which in expectation should be the same for all genes in the genome

under neutrality (Rohlfs & Nielsen, 2015). For genes that are affected by natural selection, the ratio will be increased or decreased depending on the directionality and modality of selection, much similar to the HKA test commonly used in genetic neutrality test (Hudson et al., 1987). Since the EVE model includes phylogenetic information, it can thus test different comparative hypotheses by selectively constraining parameters, including lineage-specific expression variance (as may be proxy of recent relaxed or increased selection on expression level), as well as lineage-specific shifts in constrained expression level, by taking into account within-species variance (Rohlfs & Nielsen, 2015). We think this approach could be an important supplement for the sequence-based evolutionary analyses in amphibians.

Investigating epigenetic changes during amphibians' adaptive evolution

There are diverse epigenetic mechanisms, like DNA methylation, histone modifications, and non-coding RNA activity, all of which can alter a given genotype's influence on an organism's phenotype without altering the underlying DNA bases. Through increasing, decreasing, or silencing the activities of genes, epigenetic changes can lead to novel phenotypic variations, which can be passed between generations and have the potential to contribute to adaptations (Richards, 2006). Model analysis further found that epigenetic mutations can speed adaptations or lead to populations with higher fitness, with small-effect epigenetic mutations generally speeding adaptations (Kronholm & Collins, 2016). These results provide an opportunity to examine whether epigenetic changes play a role in the adaptive evolution of amphibians, although the extent that epigenetic variation contributes to phenotypic variation remains largely unknown, and determining the drivers of divergence remains challenging (Hawes et al., 2018).

Current epigenetics studies are focused primarily on DNA methylation, because it is the most common epigenetic mark and there have been a range of simple experimental methods to detect and quantify it (Laird, 2010; Plongthongkum et al., 2014). As expect, extensive studies evidenced that DNA methylation may play a key role in rapid adaptation to environmental stress (e.g., thermotolerance), particularly in the absence of genetic variation (Dai et al., 2017; Hawes et al., 2018; Richards et al., 2012). A complex framework has been formed in term of genetic–epigenetic-environment interactions in studies of ecological epigenetics, which can also be used to study the contributions of epigenetic changes in adaption (Artemov et al., 2017; Bossdorf et al., 2010; Dubin et al., 2015; Herrera & Bazaga, 2011).

Amphibians serve as an excellent system to study whether the epigenetic variations contribute to their adaptive evolution, particularly because there are many ecologically-generalist species that inhabit a wide environmental gradient and present observable phenotypic differences between populations (i.e., *N. parkeri*). This is a prerequisite for the study of genetic–epigenetic-environment interactions (Hawes

et al., 2018). Since amphibians are not clonal species, which can result in genetically identical individuals, one problem needs to address is to distinguish between the effects of genetic variation and epigenetic changes, for which both common garden and reciprocal transplant experiments could be useful (Richards et al., 2012). With the wide application of whole-genome bisulfite sequencing (WGBS), investigating the genes and regulatory regions with different methylation during amphibians' local adaptation would be another research focus and trends in the near future.

Investigating dynamic evolution of transposable elements in amphibians

Transposable elements (TEs) are known as diverse mobile sequences that have proliferated extensively throughout eukaryotic genomes. There have been accumulating evidences supporting that transposable elements are powerful drivers of genome evolution, whose activities can considerably impact genome structures and functions, through gene disruption, mediating genomic rearrangements that cause translocation, duplication or deletion of genetic material, or affecting proximal gene expression (Fedoroff, 2012; Seidl & Thomma, 2017). Although the genetic variability induced by TEs are generally purged from the population by purifying selection, similar to other genetic variations, TE-induced mutations might also contribute to adaptive evolution by increasing the gene function or expression.

There are three main ways through which TEs can affect the evolution of the host genes and play a role in the host adaptations. The first one refers to "molecular domestication", during which some TE proteins can be recruited into the host gene interaction networks with some modifications and then play different but important roles for host survival (Miller et al., 2000; Sinzelle et al., 2009). One standard example is the evolution of RAG1–RAG2 recombinase (RAG) which originates from the ancestral RAG transposase with some adaptive amino acid changes (Zhang et al., 2019).

The second one is to change the gene structure by inserting into introns, which leads to their activation as alternatively spliced cassette exons, an event called exonization (Sorek, 2007). In this process, TEs may generate novel exons or introns for the host genes and become an important contributor to evolution and speciation (Sela et al., 2010). It was found that there is a large fraction of the TEs inserted into transcribed intronic regions of both human and mouse (Sela et al., 2007), providing the possibility for TE playing a role in the adaptive evolution of species. Further, exonizations can be population-specific and thus could contribute on population divergence (Sela et al., 2010). For example, Rech et al., (2019) screened 303 *Drosophila melanogaster* genomes from 60 worldwide natural populations and identified a set of candidate adaptive TE insertions that are present at high population frequencies. Genes located nearby these sets of TE insertions were found significantly enriched for previously identified loci underlying stress- and behavior-related traits, such as responses to stimulus, behaviors, and developments,

which give the evidence for the widespread contribution of TEs to adaptive evolution (Rech et al., 2019).

The third way in which TE may contribute to adaptive evolution is to provide regulatory elements (e.g., promoter) for the nearby genes and then improve the ability of transcriptional response of species to a common set of external stresses and environmental shock (Faulkner et al., 2009; Melé et al., 2015). For example, some people found that TE mobility in fission yeast was greatly increased when cells were exposed to unusual forms of stress (i.e., heavy metals, caffeine, and the plasticizer phthalate), and the novel TE insertions can provide the major path to resistance by linking to TOR (target of rapamycin) regulation and metal response genes (Esnault et al., 2019).

A typical feature of amphibians is the abundance of TEs within their genomes (http://www.genomesize.com). Recently, it was found that amphibians (i.e., *Oophaga pumilio*) had acquired some horizontally transferred TEs from other taxa, and some of these elements are present in a high number of copy with up-regulated expression levels, suggesting their high activity and ongoing proliferation (Rogers et al., 2018). However, whether and how TEs play roles in the adaptive evolution of amphibians remains unclear. For this question, an important prerequisite is to have a high-quality genome sequence, which can be used as a reference to conduct population genomic analyses so as to explorer the dynamic evolution of TEs during the evolution of the organism. With the increasing availability of amphibian genomes, the functional roles of TEs in adaptive evolution might become a new research hotspot.

CONCLUSIONS

Amphibians are an organismal group with extremely rich species diversity and remarkable evolutionary innovations regarding the transition from aquatic to terrestrial lifestyle. This group also provides rich bioactive substances that can be applicable to human health, particularly the skin antioxidant peptides (Mwangi et al., 2019). As such, amphibians provide abundant resources for studying the evolutionary and genetic mechanisms underlying particular phenotypes, one important topic of the evolutionary biology. However, even with the fast developments in genomics, the genomic study of amphibians has just started, partly due to the unique challenges such as large genome size and rich transposable elements that this organismal group present. In the next few years, the study on the adaptive evolution of amphibians will focus on two main aspects, namely (1) the accumulation of more high-quality reference genomes through the newly developed sequencing techniques and assembly methods, and (2) developments and implementations of new analytical strategies and methods in the amphibian systems. Through this review, we hope that our perspectives listed here provide some valuable references for evolutionary herpetologists to designate their research plans, particularly on the bioinformatics analyses. We also hope that this review can invoke discussions and attentions on the methodological aspects of adaptive evolution studies in general, particularly regarding changing from the outlier-based methods to the network-based ones and the involvement of positive selection detection based on expression matrix.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Y.B.S. conceived the review. Y.B.S., Y.Z., and K.W. prepared the draft. All authors contributed to the discussions. All authors read and approved the final version of the manuscript.

REFERENCES

Allendorf FW. 2017. Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology*, **26**(2): 420−430.

Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. *Nature Reviews Genetics*, **11**(10): 697−709.

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, **17**(2): 81−92.

Artemov AV, Mugue NS, Rastorguev SM, Zhenilo S, Mazur AM, Tsygankova SV, Boulygina ES, Kaplun D, Nedoluzhko AV, Medvedeva YA, Prokhortchouk EB. 2017. Genome-wide DNA methylation profiling reveals epigenetic adaptation of stickleback to marine and freshwater conditions. *Molecular Biology and Evolution*, **34**(9): 2203−2213.

Beall CM, Cavalleri GL, Deng LB, Elston RC, Gao Y, Knight J, Li CH, Li JC, Liang Y, McCormack M, Montgomery HE, Pan H, Robbins PA, Shianna KV, Tam SC, Tsering N, Veeramah KR, Wang W, Wangdui P, Weale ME, Xu YM, Xu Z, Yang L, Zaman MJ, Zeng CQ, Zhang L, Zhang XL, Zhaxi PC, Zheng YT. 2010. Natural selection on *EPAS1* (*HIF2α*) associated with low hemoglobin concentration in Tibetan highlanders. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(25): 11459−11464.

Beaumont MA, Balding DJ. 2004. Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**(4): 969−980.

Böcher TW. 1977. Convergence as an evolutionary process. *Botanical Journal of the Linnean Society*, **75**(1): 1−19.

Bossdorf O, Arcuri D, Richards CL, Pigliucci M. 2010. Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in *Arabidopsis thaliana*. *Evolutionary Ecology*, **24**(3): 541−553.

Brandies P, Peel E, Hogg CJ, Belov K. 2019. The value of reference genomes in the conservation of threatened species. *Genes*, **10**(11): 846.

Che J, Wang K. 2016. AmphibiaChina: an online database of Chinese Amphibians. *Zoological Research*, **37**(1): 57−59.

Chen J, Swofford R, Johnson J, Cummings BB, Rogel N, Lindblad-Toh K, Haerty W, Di Palma F, Regev A. 2019. A quantitative framework for characterizing the evolutionary history of mammalian gene expression. *Genome Research*, **29**(1): 53−63.

Chikina M, Robinson JD, Clark NL. 2016. Hundreds of genes experienced convergent shifts in selective pressure in marine mammals. *Molecular*

Biology and Evolution, **33**(9): 2182−2192.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK. 2010. Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**(4): 1411−1423.

Dai TM, Lü ZC, Liu WX, Wan FH, Hong XY. 2017. The homology gene *BtDnmt1* is essential for temperature tolerance in invasive *Bemisia* tabaci Mediterranean cryptic species. *Scientific Reports*, **7**(1): 3040.

Daub JT, Hofer T, Cutivet E, Dupanloup I, Quintana-Murci L, Robinson-Rechavi M, Excoffier L. 2013. Evidence for polygenic adaptation to pathogens in the human genome. *Molecular Biology and Evolution*, **30**(7): 1544−1558.

Deamer D, Akeson M, Branton D. 2016. Three decades of nanopore sequencing. *Nature Biotechnology*, **34**(5): 518−524.

Denton RD, Kudra RS, Malcom JW, Preez LD, Malone JH. 2018. The African Bullfrog (*Pyxicephalus adspersus*) genome unites the two ancestral ingredients for making vertebrate sex chromosomes. *bioRxiv*, doi: 10.1101/329847.

Di Franco A, Poujol R, Baurain D, Philippe H. 2019. Evaluating the usefulness of alignment filtering methods to reduce the impact of errors on evolutionary inferences. *BMC Evolutionary Biology*, **19**(1): 21.

Dobzhansky T, Gould SJ. 1982. Genetics and the Origin of Species. Columbia: Columbia University Press.

Dubin MJ, Zhang P, Meng DZ, Remigereau MS, Osborne EJ, Casale FP, Drewe P, Kahles A, Jean G, Vilhjálmsson B, Jagoda J, Irez S, Voronin V, Song Q, Long Q, Rätsch G, Stegle O, Clark RM, Nordborg M. 2015. DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *eLife*, **4**: e05255.

Edwards RJ, Tuipulotu DE, Amos TG, O'Meally D, Richardson MF, Russell TL, Vallinoto M, Carneiro M, Ferrand N, Wilkins MR, Sequeira F, Rollins LA, Holmes EC, Shine R, White PA. 2018. Draft genome assembly of the invasive cane toad, *Rhinella marina*. *Gigascience*, **7**(9): giy095.

Elewa A, Wang H, Talavera-López C, Joven A, Brito G, Kumar A, Hameed LS, Penrad-Mobayed M, Yao ZY, Zamani N, Abbas Y, Abdullayev I, Sandberg R, Grabherr M, Andersson B, Simon A. 2017. Reading and editing the *Pleurodeles waltl* genome reveals novel features of tetrapod regeneration. *Nature Communications*, **8**(1): 2286.

Elliott TA, Gregory TR. 2015. What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **370**(1678): 20140331.

Esnault C, Lee M, Ham C, Levin HL. 2019. Transposable element insertions in fission yeast drive adaptation to environmental stress. *Genome Research*, **29**(1): 85−95.

Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, Irvine KM, Schroder K, Cloonan N, Steptoe AL, Lassmann T, Waki K, Hornig N, Arakawa T, Takahashi H, Kawai J, Forrest ARR, Suzuki H, Hayashizaki Y, Hume DA, Orlando V, Grimmond SM, Carninci P. 2009. The regulated retrotransposon transcriptome of mammalian cells. *Nature Genetics*, **41**(5): 563−571.

Fedoroff NV. 2012. Presidential address. *Transposable elements, epigenetics, and genome evolution. Science*, **338**(6108): 758−767.

Felsenstein J. 1985. Phylogenies and the comparative method. *The American Naturalist*, **125**(1): 1−15.

Ferea TL, Botstein D, Brown PO, Rosenzweig RF. 1999. Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, **96**(17): 9721−9726.

Field Y, Boyle EA, Telis N, Gao ZY, Gaulton KJ, Golan D, Yengo L, Rocheleau G, Froguel P, McCarthy MI, Pritchard JK. 2016. Detection of human adaptation during the past 2000 years. *Science*, **354**(6313): 760−764.

Fletcher W, Yang ZH. 2010. The effect of insertions, deletions, and alignment errors on the branch-site test of positive selection. *Molecular Biology and Evolution*, **27**(10): 2257−2267.

Foll M, Gaggiotti OE, Daub JT, Vatsiou A, Excoffier L. 2014. Widespread signals of convergent adaptation to high altitude in Asia and america. *The American Journal of Human Genetics*, **95**(4): 394−407.

Foote AD, Liu Y, Thomas GWC, Vinař T, Alföldi J, Deng JX, Dugan S, Van Elk CE, Hunter ME, Joshi V, Khan Z, Kovar C, Lee SL, Lindblad-Toh K, Mancia A, Nielsen R, Qin X, Qu JX, Raney BJ, Vijay N, Wolf JBW, Hahn MW, Muzny DM, Worley KC, Gilbert MTP, Gibbs RA. 2015. Convergent evolution of the genomes of marine mammals. *Nature Genetics*, **47**(3): 272−275.

Fraser HB, Moses AM, Schadt EE. 2010. Evidence for widespread adaptive evolution of gene expression in budding yeast. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(7): 2977−2982.

Funk WC, Murphy MA, Hoke KL, Muths E, Amburgey SM, Lemmon EM, Lemmon AR. 2016. Elevational speciation in action? Restricted gene flow associated with adaptive divergence across an altitudinal gradient. *Journal of Evolutionary Biology*, **29**(2): 241−252.

Funk WC, Zamudio KR, Crawford AJ. 2018. Advancing Understanding of Amphibian Evolution, Ecology, Behavior, and Conservation with Massively Parallel Sequencing. Cham: Springer.

Ge RL, Cai QL, Shen YY, San A, Ma L, Zhang Y, Yi X, Chen Y, Yang LF, Huang Y, He RJ, Hui YY, Hao MR, Li Y, Wang B, Ou XH, Xu JH, Zhang YF, Wu K, Geng CY, Zhou WP, Zhou TC, Irwin DM, Yang YZ, Ying L, Bao HH, Kim J, Larkin DM, Ma J, Lewin HA, Xing JC, Platt II RN, Ray DA, Auvil L, Capitanu B, Zhang XF, Zhang GJ, Murphy RW, Wang J, Zhang YP, Wang J. 2013. Draft genome sequence of the Tibetan antelope. *Nature Communications*, **4**: 1858.

Genome 10K Community of Scientists. 2009. Genome 10K: a proposal to obtain whole-genome sequence for 10000 vertebrate species. *Journal of Heredity*, **100**(6): 659−674.

Glenn TC. 2011. Field guide to next-generation DNA sequencers. *Molecular Ecology Resources*, **11**(5): 759−769.

Gnecchi-Ruscone GA, Abondio P, De Fanti S, Sarno S, Sherpa MG, Sherpa PT, Marinelli G, Natali L, Di Marcello M, Peluzzi D, Luiselli D, Pettener D, Sazzini M. 2018. Evidence of polygenic adaptation to high altitude from tibetan and sherpa genomes. *Genome Biology and Evolution*, **10**(11): 2919−2930.

Gouy A, Daub JT, Excoffier L. 2017. Detecting gene subnetworks under selection in biological pathways. *Nucleic Acids Research*, **45**(16): e149..

Guo BC, Lu D, Liao WB, Merila J. 2016. Genomewide scan for adaptive differentiation along altitudinal gradient in the Andrew's toad *Bufo andrewsi*. *Molecular Ecology*, **25**(16): 3884−3900.

Guo WF, Zhang SW, Zeng T, Li Y, Gao JX, Chen LN. 2019. A novel network control model for identifying personalized driver genes in cancer. *PLoS Computational Biology*, **15**(11): e1007520.

Hammond SA, Warren RL, Vandervalk BP, Kucuk E, Khan H, Gibb EA, Pandoh P, Kirk H, Zhao YJ, Jones M, Mungall AJ, Coope R, Pleasance S,

Moore RA, Holt RA, Round JM, Ohora S, Walle BV, Veldhoen N, Helbing CC, Birol I. 2017. The North American bullfrog draft genome provides insight into hormonal regulation of long noncoding RNA. *Nature Communications*, **8**(1): 1433.

Hancock AM, Witonsky DB, Ehler E, Alkorta-Aranburu G, Beall C, Gebremedhin A, Sukernik R, Utermann G, Pritchard J, Coop G, Di Rienzo A. 2010. Colloquium paper: human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(Suppl 2): 8924−8930.

Hansen TF. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution*, **51**(5): 1341−1351.

Hao Y, Xiong Y, Cheng YL, Song G, Jia CX, Qu YH, Lei FM. 2019. Comparative transcriptomics of 3 high-altitude passerine birds and their low-altitude relatives. *Proceedings of the National Academy of Sciences of the United States of America*, **116**(24): 11851−11856.

Hawes NA, Fidler AE, Tremblay LA, Pochon X, Dunphy BJ, Smith KF. 2018. Understanding the role of DNA methylation in successful biological invasions: a review. *Biological Invasions*, **20**(9): 2285−2300.

Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, Kapitonov V, Ovcharenko I, Putnam NH, Shu SQ, Taher L, Blitz IL, Blumberg B, Dichmann DS, Dubchak I, Amaya E, Detter JC, Fletcher R, Gerhard DS, Goodstein D, Graves T, Grigoriev IV, Grimwood J, Kawashima T, Lindquist E, Lucas SM, Mead PE, Mitros T, Ogino H, Ohta Y, Poliakov AV, Pollet N, Robert J, Salamov A, Sater AK, Schmutz J, Terry A, Vize PD, Warren WC, Wells D, Wills A, Wilson RK, Zimmerman LB, Zorn AM, Grainger R, Grammer T, Khokha MK, Richardson PM, Rokhsar DS. 2010. The genome of the Western clawed frog *Xenopus tropicalis*. *Science*, **328**(5978): 633−636.

Herrera CM, Bazaga P. 2011. Untangling individual variation in natural populations: ecological, genetic and epigenetic correlates of long-term inequality in herbivory. *Molecular Ecology*, **20**(8): 1675−1688.

Hudson RR, Kreitman M, Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**(1): 153−159.

Idaghdour Y, Czika W, Shianna KV, Lee SH, Visscher PM, Martin HC, Miclaus K, Jadallah SJ, Goldstein DB, Wolfinger RD, Gibson G. 2009. Geographical genomics of human leukocyte gene expression variation in southern Morocco. *Nature Genetics*, **42**(1): 62−67.

King MC, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees. *Science*, **188**(4184): 107−116.

Koepfli KP, Paten B, The Genome 10K Community of Scientists, O'brien SJ. 2015. The Genome 10K Project: a way forward. *Annual Review of Animal Biosciences*, **3**: 57−111.

Kronholm I, Collins S. 2016. Epigenetic mutations can both help and hinder adaptive evolution. *Molecular Ecology*, **25**(8): 1856−1868.

Laird PW. 2010. Principles and challenges of genomewide DNA methylation analysis. *Nature Reviews Genetics*, **11**(3): 191−203.

Lamichhaney S, Berglund J, Almen MS, Maqbool K, Grabherr M, Martinez-Barrio A, Promerová M, Rubin CJ, Wang C, Zamani N, Grant BR, Grant PR, Webster MT, Andersson L. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature*, **518**(7539): 371−375.

Lee KM, Coop G. 2019. Population genomics perspectives on convergent adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **374**(1777): 20180236.

Li G, Wang JH, Rossiter SJ, Jones G, Cotton JA, Zhang SY. 2008. The hearing gene *Prestin reunites* echolocating bats. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(37): 13959−13964.

Li J, Yu HY, Wang WX, Fu C, Zhang W, Han FM, Wu H. 2019a. Genomic and transcriptomic insights into molecular basis of sexually dimorphic nuptial spines in *Leptobrachium leishanense*. *Nature Communications*, **10**(1): 5551.

Li JT, Gao YD, Xie L, Deng C, Shi P, Guan ML, Huang S, Ren JL, Wu DD, Ding L, Huang ZY, Nie H, Humphreys DP, Hillis DM, Wang WZ, Zhang YP. 2018. Comparative genomic investigation of high-elevation adaptation in ectothermic snakes. *Proceedings of the National Academy of Sciences of the United States of America*, **115**(33): 8406−8411.

Li Y, Liu Z, Shi P, Zhang JZ. 2010. The hearing gene *Prestin* unites echolocating bats and whales. *Current Biology*, **20**(2): R55−R56.

Li YX, Ren YD, Zhang DR, Jiang H, Wang ZK, Li XY, Rao DQ. 2019b. Chromosome-level assembly of the mustache toad genome using third-generation DNA sequencing and Hi-C analysis. *Gigascience*, **8**(9): giz114.

Liang L, Shen YY, Pan XW, Zhou TC, Yang C, Irwin DM, Zhang YP. 2013. Adaptive evolution of the *Hox* gene family for development in bats and dolphins. *PLoS One*, **8**(6): e65944.

Liedtke HC, Gower DJ, Wilkinson M, Gomez-Mestre I. 2018. Macroevolutionary shift in the size of amphibian genomes and the role of life history and climate. *Nature Ecology & Evolution*, **2**(11): 1792−1799.

Löytynoja A, Goldman N. 2005. An algorithm for progressive multiple alignment of sequences with insertions. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(30): 10557−10562.

Marouli E, Graff M, Medina-Gomez C, et al. 2017. Rare and low-frequency coding variants alter human adult height. *Nature*, **542**(7640): 186−190.

Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ, Johnson R, Segrè AV, Djebali S, Niarchou A, The GTEx Consortium, Wright FA, Lappalainen T, Calvo M, Getz G, Dermitzakis ET, Ardlie KG, Guigó R. 2015. The human transcriptome across tissues and individuals. *Science*, **348**(6235): 660−665.

Miller WJ, McDonald JF, Nouaud D, Anxolabéhère D. 2000. Molecular domestication—more than a sporadic episode in evolution. *In*: McDonald JF. Transposable Elements and Genome Evolution. Dordrecht: Springer, 197−207.

Mwangi J, Hao X, Lai R, Zhang ZY. 2019. Antimicrobial peptides: new hope in the war against multidrug resistance. *Zoological Research*, **40**(6): 488−505.

Nadeau NJ, Jiggins CD. 2010. A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics*, **26**(11): 484−492.

Nowoshilow S, Schloissnig S, Fei JF, Dahl A, Pang AWC, Pippel M, Winkler S, Hastie AR, Young G, Roscito JG, Falcon F, Knapp D, Powell S, Cruz A, Cao H, Habermann B, Hiller M, Tanaka EM, Myers EW. 2018. The axolotl genome and the evolution of key tissue formation regulators. *Nature*, **554**(7690): 50−55.

Orr HA. 1998. Testing natural selection *vs*. genetic drift in phenotypic evolution using quantitative trait locus data. *Genetics*, **149**(4): 2099−2104.

Ostrander EA, Wang GD, Larson G, Vonholdt BM, Davis BW, Jagannathan V, Hitte C, Wayne RK, Zhang YP, Dog10K Consortium. 2019. Dog10K: an

international sequencing effort to advance studies of canine domestication, phenotypes and health. *National Science Review*, **6**(4): 810−824.

Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, Veyrieras JB, Stephens M, Gilad Y, Pritchard JK. 2010. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature*, **464**(7289): 768−772.

Plongthongkum N, Diep DH, Zhang K. 2014. Advances in the profiling of DNA modifications: cytosine methylation and beyond. *Nature Reviews Genetics*, **15**(10): 647−661.

Prohaska A, Racimo F, Schork AJ, Sikora M, Stern AJ, Ilardo M, Allentoft ME, Folkersen L, Buil A, Moreno-Mayar JV, Korneliussen T, Geschwind D, Ingason A, Werge T, Nielsen R, Willerslev E. 2019. Human disease variation in the light of population genomics. *Cell*, **177**(1): 115−131.

Rech GE, Bogaerts-Márquez M, Barrón MG, Merenciano M, Villanueva-Cañas JL, Horváth V, Fiston-Lavier AS, Luyten I, Venkataram S, Quesneville H, Petrov DA, González J. 2019. Stress response, behavior, and development are shaped by transposable element-induced mutations in Drosophila. *PLoS Genetics*, **15**(2): e1007900.

Richards CL, Schrey AW, Pigliucci M. 2012. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. *Ecology Letters*, **15**(9): 1016−1025.

Richards EJ. 2006. Inherited epigenetic variation--revisiting soft inheritance. *Nature Reviews Genetics*, **7**(5): 395−401.

Rogers RL, Zhou L, Chu C, Márquez R, Corl A, Linderoth T, Freeborn L, MacManes MD, Xiong ZJ, Zheng J, Guo CX, Xun X, Kronforst MR, Summers K, Wu YF, Yang HM, Richards-Zawacki CL, Zhang GJ, Nielsen R. 2018. Genomic takeover by transposable elements in the strawberry poison frog. *Molecular Biology and Evolution*, **35**(12): 2913−2927.

Rohlf FJ. 2001. Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution*, **55**(11): 2143−2160.

Rohlfs RV, Nielsen R. 2015. Phylogenetic ANOVA: the expression variance and evolution model for quantitative trait evolution. *Systematic Biology*, **64**(5): 695−708.

Rollins-Smith LA. 2009. The role of amphibian antimicrobial peptides in protection of amphibians from pathogens linked to global amphibian declines. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, **1788**(8): 1593−1599.

Ruan J, Li H. 2020. Fast and accurate long-read assembly with wtdbg2. *Nature Methods*, **17**(2): 155−158.

Seidl F, Levis NA, Schell R, Pfennig DW, Pfennig KS, Ehrenreich IM. 2019. Genome of *Spea multiplicata*, a rapidly developing, phenotypically plastic, and desert-adapted spadefoot toad. *G3: Genes, Genomes, Genetics*, **9**(12): 3909−3919.

Seidl MF, Thomma BPHJ. 2017. Transposable elements direct the coevolution between plants and microbes. *Trends in Genetics*, **33**(11): 842−851.

Sela N, Mersch B, Gal-Mark N, Lev-Maor G, Hotz-Wagenblatt A, Ast G. 2007. Comparative analysis of transposed element insertion within human and mouse genomes reveals *Alu's* unique role in shaping the human transcriptome. *Genome Biology*, **8**(6): R127.

Sela N, Mersch B, Hotz-Wagenblatt A, Ast G. 2010. Characteristics of transposable element exonization within human and mouse. *PLoS One*, **5**(6): e10907.

Session AM, Uno Y, Kwon T, et al. 2016. Genome evolution in the allotetraploid frog Xenopus laevis. *Nature*, **538**(7625): 336−343.

Shaffer HB, Gidis M, McCartney-Melstad E, Neal KM, Oyamaguchi HM, Tellez M, Toffelmier EM. 2015. Conservation genetics and genomics of amphibians and reptiles. *Annual Review of Animal Biosciences*, **3**: 113−138.

Shen JX, Feng AS, Xu ZM, Yu ZL, Arch VS, Yu XJ, Narins PM. 2008. Ultrasonic frogs show hyperacute phonotaxis to female courtship calls. *Nature*, **453**(7197): 914−916.

Sinzelle L, Izsvák Z, Ivics Z. 2009. Molecular domestication of transposable elements: from detrimental parasites to useful host genes. *Cellular and Molecular Life Sciences*, **66**(6): 1073−1093.

Smith JJ, Timoshevskaya N, Timoshevskiy VA, Keinath MC, Hardy D, Voss SR. 2019. A chromosome-scale assembly of the axolotl genome. *Genome Research*, **29**(2): 317−324.

Solak HM, Yanchukov A, Çolak F, Matur F, Sözen M, Ayanoğlu İC, Winternitz JC. 2020. Altitudinal Effects on Innate Immune Response of a Subterranean Rodent. *Zoological Science*, **37**(1): 31−41.

Sorek R. 2007. The birth of new exons: mechanisms and evolutionary consequences. *RNA*, **13**(10): 1603−1608.

Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Research*, **32**(S2): W309-−W312..

Storfer A, Eastman JM, Spear SF. 2009. Modern molecular methods for amphibian conservation. *BioScience*, **59**(7): 559−571.

Sun YB. 2017. FasParser: a package for manipulating sequence data. *Zoological Research*, **38**(2): 110−112.

Sun YB. 2018. FasParser2: a graphical platform for batch manipulation of tremendous amount of sequence data. *Bioinformatics*, **34**(14): 2493−2495.

Sun YB, Fu TT, Jin JQ, Murphy RW, Hillis DM, Zhang YP, Che J. 2018. Species groups distributed across elevational gradients reveal convergent and continuous genetic adaptation to high elevations. *Proceedings of the National Academy of Sciences of the United States of America*, **115**(45): E10634−E10641.

Sun YB, Xiong ZJ, Xiang XY, Liu SP, Zhou WW, Tu XL, Zhong L, Wang L, Wu DD, Zhang BL, Zhu CL, Yang MM, Chen HM, Li F, Zhou L, Feng SH, Huang C, Zhang GJ, Irwin D, Hillis DM, Murphy RW, Yang HM, Che J, Wang J, Zhang YP. 2015. Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(11): E1257−E1262.

Supple MA, Shapiro B. 2018. Conservation of biodiversity in the genomics era. *Genome Biology*, **19**(1): 131.

The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**(7414): 57−74.

Thomas GWC, Hahn MW. 2015. Determining the null model for detecting adaptive convergence from genomic data: a case study using echolocating mammals. *Molecular Biology and Evolution*, **32**(5): 1232−1236.

Tian H, Hammer RE, Matsumoto AM, Russell DW, McKnight SL. 1998. The hypoxia-responsive transcription factor EPAS1 is essential for catecholamine homeostasis and protection against heart failure during embryonic development. *Genes & Development*, **12**(21): 3320−3324.

Vitti JJ, Grossman SR, Sabeti PC. 2013. Detecting natural selection in genomic data. *Annual Review of Genetics*, **47**: 97−120.

Wang B, Chen L, Wang W. 2019. Genomic insights into ruminant evolution:

from past to future prospects. *Zoological Research*, **40**(6): 476−487.

Wang GD, Zhang BL, Zhou WW, Li YX, Jin JQ, Shao Y, Yang HC, Liu YH, Yan F, Chen HM, Jin L, Gao F, Zhang YG, Li HP, Mao BY, Murphy RW, Wake DB, Zhang YP, Che J. 2018. Selection and environmental adaptation along a path to speciation in the Tibetan frog *Nanorana parkeri*. *Proceedings of the National Academy of Sciences of the United States of America*, **115**(22): E5056−E5065.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, **10**(1): 57−63.

Ward RD, Skibinski DOF, Woodwark M. 1992. Protein heterozygosity, protein structure, and taxonomic differentiation. *In*: Hecht MK, Wallace B, Macintyre RJ. Evolutionary Biology. Boston: Springer, 73–159.

Whelan S, Irisarri I, Burki F. 2018. PREQUAL: detecting non-homologous characters in sets of unaligned homologous sequences. *Bioinformatics*, **34**(22): 3929−3930.

Wood AR, Esko T, Yang J, et al. 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, **46**(11): 1173−1186.

Yang CH, Fu TT, Lan XQ, Zhang Y, Nneji LM, Murphy RW, Sun YB, Che J. 2019. Comparative skin histology of frogs reveals high-elevation adaptation of the tibetan *Nanorana parkeri*. *Asian Herpetological Research*, **10**(2): 79−85.

Yang JR, Maclean CJ, Park C, Zhao HB, Zhang JZ. 2017. Intra and interspecific variations of gene expression levels in yeast are largely neutral: (Nei lecture, SMBE 2016, gold coast). *Molecular Biology and Evolution*, **34**(9): 2125−2139.

Yang WZ, Qi Y, Bi K, Fu JZ. 2012. Toward understanding the genetic basis of adaptation to high-elevation life in poikilothermic species: a comparative transcriptomic analysis of two ranid frogs, *Rana chensinensis* and *R. kukunoris*. *BMC Genomics*, **13**: 588.

Yang XW, Wang Y, Zhang Y, Lee WH, Zhang Y. 2016. Rich diversity and potency of skin antioxidant peptides revealed a novel molecular basis for high-altitude adaptation of amphibians. *Scientific Reports*, **6**: 19866.

Yu XT, Zhang JS, Sun SY, Zhou X, Zeng T, Chen LN. 2017. Individual-specific edge-network analysis for disease prediction. *Nucleic Acids Research*, **45**(20): e170.

Zhang GJ, Li C, Li QY, et al. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, **346**(6215): 1311−1320.

Zhang GJ. 2015. Genomics: bird sequencing project takes off. *Nature*, **522**(7554): 34.

Zhang JZ, Nielsen R, Yang ZH. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular Biology and Evolution*, **22**(12): 2472−2479.

Zhang YH, Cheng TC, Huang GR, Lu QY, Surleac MD, Mandell JD, Pontarotti P, Petrescu AJ, Xu AL, Xiong Y, Schatz DG. 2019. Transposon molecular domestication and the evolution of the RAG recombinase. *Nature*, **569**(7754): 79−84.