

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/256077642>

Genomes and evolutionary genomics of animals

Article in *Current Zoology* · February 2013

DOI: 10.1093/czoolo/59.1.87

CITATIONS

3

READS

115

2 authors, including:



Luting Song

Beijing Genomics Institute

7 PUBLICATIONS 961 CITATIONS

SEE PROFILE

Genomes and evolutionary genomics of animals

Luting SONG^{1,2}, Wen WANG^{1*}

¹ State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

² Graduate University of Chinese Academy of Sciences, Beijing 100049, China

Abstract Alongside recent advances and booming applications of DNA sequencing technologies, a great number of complete genome sequences for animal species are available to researchers. Hundreds of animals have been involved in whole genome sequencing, and at least 87 non-human animal species' complete or draft genome sequences have been published since 1998. Based on these technological advances and the subsequent accumulation of large quantity of genomic data, evolutionary genomics has become one of the most rapidly advancing disciplines in biology. Scientists now can perform a number of comparative and evolutionary genomic studies for animals, to identify conserved genes or other functional elements among species, genomic elements that confer animals their own specific characteristics and new phenotypes for adaptation. This review deals with the current genomic and evolutionary research on non-human animals, and displays a comprehensive landscape of genomes and the evolutionary genomics of non-human animals. It is very helpful to a better understanding of the biology and evolution of the myriad forms within the animal kingdom [*Current Zoology* 59 (1): 87–98, 2013].

Keywords Animals, Genome, Evolutionary genomics

1 Introduction

Since the *Cambrian* explosion some 530 million years ago (mya), evolution in multi-cellular animals has been an ongoing process, resulting in the tremendous biodiversity of the animal kingdom. At the core of evolution is the change in DNA sequences. Accordingly, for the evolutionary study of animals, it is imperative to investigate changes in the genome sequences. Thanks in large part to the Human Genome Project, formally began in 1990, a great number of genome sequencing technologies and bioinformatic tools have greatly extended our knowledge of genomics and thus life science. With the booming applications of next-generation DNA sequencing technologies that have a noticeably decreased cost from several years ago, it is possible for numerous researchers to complete genome sequences of animals for use in evolutionary genomic studies. To coincide with these developments, majestic genome projects that aim to decipher the most phylogenetically and economically important species in kingdom *Animalia* have emerged, such as BGI-1,000 Plant and Animal Reference Genomes Project, Genome 10K Project, i5k Insect and other Arthropod Genome Sequencing Initiative, and NIH National Human Genome Research Institute -Approved Sequencing Targets as well.

Given the massive sequence information, a great deal of insightful comparative evolutionary analyses across these animals and across the biological and academic spectrums are expected. Indeed, evolutionary genomics is quickly becoming one of the most rapidly advancing disciplines in biological sciences. To present a comprehensive landscape of genomics and in particular the evolutionary genomics of non-human animals with published papers on their genome projects, after giving a list of non-human animals with publications on their genome projects, we shall first review genomic studies for model animals, and then introduce such studies for other non-human animals following an evolutionary hierarchy of animal phyla. These studies touch on conserved genes or other functional elements among non-human animal species, as well as genomic elements that confer an organism its own specific characteristics and phenotypes for adaptation. In doing so, we hope to clearly illustrate the marked gains and insights that evolutionary genomics have already made and still have to offer.

2 Animal Species Involved in Genome Sequencing

To mine completely sequenced genomes of non-

Received Oct. 18, 2012; accepted Nov. 8, 2012.

* Corresponding author. E-mail: wwang@mail.kiz.ac.cn

© 2013 *Current Zoology*

human animals, we searched the relevant literatures and such main genome projects and databases as BGI-1,000 Plant and Animal Reference Genomes Project (<http://ldl.genomics.cn/page/pa-animal.jsp>), Genome 10K Project (<http://www.genome10k.org/>), i5k Insect and other Arthropod Genome Sequencing Initiative (<http://www.arthropodgenomes.org/wiki/i5K>), NIH National Human Genome Research Institute-Approved Sequencing Targets (<http://www.genome.gov/10002154>), and the database Ensembl Metazoa Species (<http://asia.ensembl.org/info/about/species.html>; <http://metazoa.ensembl.org/info/about/species.html>).

We located over one hundred animals that have been sequencing-completed, including house mouse, goat, giant panda, and so on; as well as thousands of animals whose whole genome sequencing is in-progress or proposed as of June 2012. Among them, at least 87 non-human animal species' complete or draft genome sequences have published papers on their genome projects since 1998 (Table 1). Also, certain other animals have released their genome sequences without genome-project publications, such as zebrafish *Danio rerio* (http://www.sanger.ac.uk/Projects/D_rerio/), domestic goat *Capra hircus* (<http://goat.kiz.ac.cn/>) and so forth. Insects are the taxon with the most sequenced-species, followed with mammals whose sequences consist 25% of all animal species' genomes (Fig. 1). In analyzing the temporal trends as to the number of publications on non-human animal whole genome sequences since 1998, we found that studies deciphering animal genomes have continually increased. Notably, this trend was all the more obvious after we excluded 10 fruit fly genomes that were published collectively in a single paper (Fig. 2).

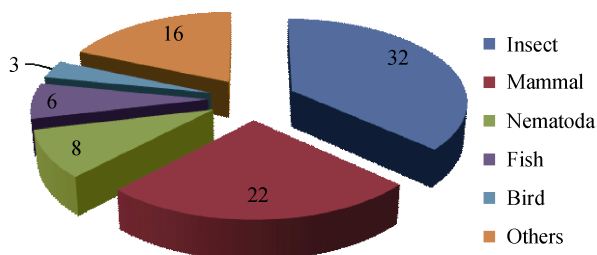


Fig. 1 Pie chart of different non-human animal taxa with publications on their genome projects

3 Model Animals

3.1 Nematodes

The phylum *Nematoda* includes not only the important model organism *Caenorhabditis elegans* but several global parasites of humans, food animals and crops.

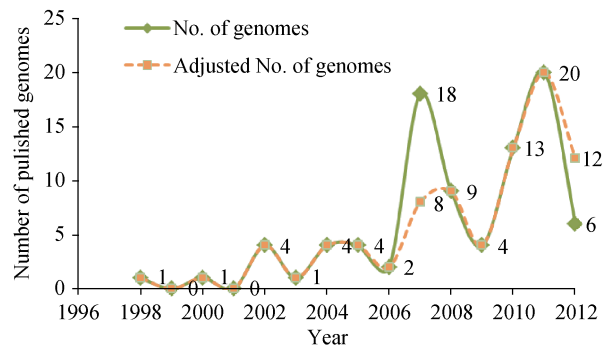


Fig. 2 Temporal trends in the number of publications on animal whole genome project between 1998 and June 2012

The solid line was generated by observed raw data, and dash line is an adjustment of 1): excluding an unusual event (10 fruit fly genomes together published in a single paper); and 2): proportional increase of value of 2012 as a whole year from that of June of 2012.

Since completion of the first animal genome sequence for *C. elegans* (*C. elegans* Sequencing Consortium 1998), draft genome sequences have been published for another seven worms from lineages of the pan-phylum *Nematoda*, including two free-living worms *C. briggsae* (Stein et al., 2003) and *Pristionchus pacificus* (Dieterich et al., 2008), two plant parasites *Meloidogyne incognita* (Abad et al., 2008) and *M. hapla* (Opperman et al., 2008), and three animal/human parasites *Trichinella spiralis* (Mitreva et al., 2011), *Brugia malayi* (Ghedini et al., 2007), and *Ascaris suum* (Jex et al., 2011). Their genome size ranges significantly from 54 Mb of *M. hapla* to 272 Mb of *A. suum*, and the gene number varies from 11,500 of *B. malayi* to 23,500 of *P. pacificus*. Overall genome organization is also not conserved among these nematodes with divergence of over 500 million year (myr). Comparative genomic analyses, however, have revealed shared genes and proteins of fundamental importance among all nematode species and their common primary biological function suggested with genomic regions of microsynteny.

Additionally, such genetic factors underlying lineage diversification have also been identified as disproportionate numbers of gene family deaths over births in parasites as compared with non-parasitic nematodes (Jex et al., 2011; Mitreva et al., 2011), and important expanding gene families among parasitic nematodes (Ghedini et al., 2007; Abad et al., 2008; Opperman et al., 2008; Jex et al., 2011; Mitreva et al., 2011).

3.2 Fruit flies

The fruit fly *Drosophila melanogaster* is among the most intensively studied organisms in biology, serving as a model system for investigating numerous genetic and developmental processes common to higher eu-

Table 1 List of non-human animals with publications on their genome projects

Scientific name	Common name	Taxonomy	Genome size Mb	Authors	Publication year
<i>Caenorhabditis elegans</i>	Soil nematode	Nematoda	100	<i>C. elegans</i> Sequencing Consortium	1998
<i>Drosophila melanogaster</i>	Fruitfly	Insect	210	Adams et al.	2000
<i>Ciona intestinalis</i>	Vase Tunicate	Ascidiacea	150	Dehal et al.	2002
<i>Fugu rubripes</i>	Globefish	Fish	400	Aparicio et al.	2002
<i>Anopheles gambiae</i>	African Malaria Mosquito	Insect	278	Holt et al.	2002
<i>Mus musculus</i>	Mouse	Mammal	2,500	Chinwalla et al.	2002
<i>Caenorhabditis briggsae</i>	Soil nematode	Nematoda	104	Stein et al.	2003
<i>Gallus gallus</i>	Red Jungle Fowl	Bird	1,200	Hillier et al.	2004
<i>Tetraodon nigroviridis</i>	Green Spotted Puffer	Fish	340	Jaillon et al.	2004
<i>Bombyx mori</i>	Silkworm	Insect	530	Xia et al.	2004
<i>Rattus norvegicus</i>	Rat	Mammal	2,500	Rat Genome Sequencing Project Consortium	2004
<i>Drosophila pseudoobscura</i>	Fruitfly	Insect	139	Richards et al.	2005
<i>Trypanosoma cruzi</i>	Trypanosome	Kinetoplastida	67	El-Sayed et al.	2005
<i>Canis familiaris</i>	Domestic Dog	Mammal	2,400	Lindblad-Toh et al.	2005
<i>Pan troglodytes</i>	Chimpanzee	Mammal	3,000	Mikkelsen et al.	2005
<i>Strongylocentrotus purpuratus</i>	Sea Urchin	Echinoidea	890	Sodergren et al.	2006
<i>Apis mellifera</i>	Honey Bee	Insect	236	The Honeybee Genome Sequencing Consortium	2006
<i>Nematostella vectensis</i>	Sea Anemone	Anthozoa	450	Putnam et al.	2007
<i>Callorhynchus milii</i>	Elephant Shark	Fish	910	Venkatesh et al.	2007
<i>Oryzias latipes</i>	Medaka	Fish	700	Kasahara et al.	2007
<i>Aedes aegypti</i>	N/A	Insect	1,310	Nene et al.	2007
<i>Drosophila ananassae</i>	Fruitfly	Insect	231	Clark et al.	2007
<i>Drosophila erecta</i>	Fruitfly	Insect	153	Clark et al.	2007
<i>Drosophila grimshawi</i>	Fruitfly	Insect	201	Clark et al.	2007
<i>Drosophila mojavensis</i>	Fruitfly	Insect	194	Clark et al.	2007
<i>Drosophila persimilis</i>	Fruitfly	Insect	189	Clark et al.	2007
<i>Drosophila sechellia</i>	Fruitfly	Insect	167	Clark et al.	2007
<i>Drosophila simulans</i>	Fruitfly	Insect	138	Clark et al.	2007
<i>Drosophila virilis</i>	Fruitfly	Insect	206	Clark et al.	2007
<i>Drosophila willistoni</i>	Fruitfly	Insect	236	Clark et al.	2007
<i>Drosophila yakuba</i>	Fruitfly	Insect	166	Clark et al.	2007
<i>Felis catus</i>	Domestic Cat	Mammal	2,700	Clark et al.	2007
<i>Macaca mulatta</i>	Rhesus Macaque	Mammal	2,870	Rhesus Macaque Genome Sequencing and Analysis Consortium	2007
<i>Monodelphis domestica</i>	Gray Short-tailed Opossum	Mammal	3,600	Mikkelsen et al.	2007
<i>Brugia malayi</i>	N/A	Nematoda	90	Ghedini et al.	2007
<i>Tribolium castaneum</i>	Red flour beetle	Insect	160	Richards et al.	2008
<i>Branchiostoma floridae</i>	Lancelet	Leptocardii	520	Putnam et al.	2008
<i>Ornithorhynchus anatinus</i>	Platypus	Mammal	2,000	Warren et al.	2008
<i>Mammuthus primigenius</i>	Woolly mammoth	Mammal	4,700	Miller et al.	2008
<i>Pristionchus pacificus</i>	Necromenic nematode	Nematoda	169	Dieterich et al.	2008
<i>Meloidogyne hapla</i>	N/A	Nematoda	54	Opperman et al.	2008
<i>Meloidogyne incognita</i>	Root-knot nematode	Nematoda	86	Abad et al.	2008
<i>Trichoplax adhaerens</i>	Trichoplax	Trichoplacoidea	98	Srivastava et al.	2008
<i>Equus caballus</i>	Domestic horse	Mammal	2,100	Wade et al.	2009

To be continued

Continued Table 1

Scientific name	Common name	Taxonomy	Genome size Mb	Authors	Publication year
<i>Bos Taurus</i>	Cattle	Mammal	3,200	Elsik et al.	2009
<i>Schistosoma mansoni</i>	Blood fluke	Trematoda	363	Berriman et al.	2009
<i>Schistosoma japonicum</i>	Blood fluke	Trematoda	397	The <i>Schistosoma japonicum</i> Genome Sequencing and Functional Analysis Consortium	2009
<i>Ailuropoda melanoleuca</i>	Giant panda	Mammal	2,400	Li et al.	2010
<i>Xenopus tropicalis</i>	Western clawed frog	Amphibia	1,700	Hellsten et al.	2010
<i>Meleagris gallopavo</i>	Domestic turkey	Bird	1,100	Dalloul et al.	2010
<i>Taeniopygia guttata</i>	Zebra finch	Bird	1,200	Warren et al.	2010
<i>Amphimedon queenslandica</i>	Sponge	Demospongiae	167	Srivastava et al.	2010
<i>Hydra magnipapillata</i>	Hydra	Hydrozoa	1,500	Chapman et al.	2010
<i>Pediculus humanus</i>	Human Body Louse	Insect	108	Kirkness et al.	2010
<i>Harpegnathos saltator</i>	Jumping Ant	Insect	240	Bonasio et al.	2010
<i>Camponotus floridanus</i>	N/A	Insect	330	Bonasio et al.	2010
<i>Nasonia vitripennis</i>	Nasonia	Insect	340	Werren et al.	2010
<i>Nasonia giraulti</i>	Nasonia	Insect	341	Werren et al.	2010
<i>Nasonia longicornis</i>	Nasonia	Insect	342	Werren et al.	2010
<i>Acyrtosiphon pisum</i>	Pea aphid	Insect	517	The International Aphid Genomics Consortium	2010
<i>Culex quinquefasciatus</i>	Southern house mosquito	Insect	540	Arensburger et al.	2010
<i>Acropora digitifera</i>	Coral	Anthozoa	420	Shinzato et al.	2011
<i>Tetranychus urticae</i>	Red spider mite	Arachnida	90	Grbic et al.	2011
<i>Daphnia pulex</i>	Water flea	Crustacea	200	Colbourne et al.	2011
<i>Gadus morhua</i>	Atlantic cod	Fish	830	Star et al.	2011
<i>Linepithema humile</i>	Argentine ant	Insect	250	Smith et al.	2011
<i>Danaus plexippus</i>	Monarch butterfly	Insect	273	Zhan et al.	2011
<i>Pogonomyrmex barbatus</i>	Red harvester ant	Insect	284	Smith et al.	2011
<i>Atta cephalotes</i>	Leafcutter ant	Insect	300	Suen et al.	2011
<i>Acromyrmex echinator</i>	Leaf-cutting ant	Insect	313	Nygaard et al.	2011
<i>Solenopsis invicta</i>	Fire ant	Insect	484	Wurm et al.	2011
<i>Macropus eugenii</i>	Kangaroo	Mammal	2,700	Renfree et al.	2011
<i>Cricetulus griseus</i>	Chinese hamster	Mammal	2,450	Xu et al.	2011
<i>Heterocephalus glaber</i>	Naked mole rat	Mammal	2,600	Kim et al.	2011
<i>Macaca mulatta lasiotea</i>	Rhesus monkey	Mammal	2,850	Yan et al.	2011
<i>Macaca fascicularis</i>	Crab-eating macaque	Mammal	2,850	Yan et al.	2011
<i>Sarcophilus harrisii</i>	Tasmanian devil	Mammal	3,300	Miller et al.	2011
<i>Pongo abelii</i>	Orang utan	Mammal	3,090	Locke et al.	2011
<i>Ascaris suum</i>	Pig roundworm	Nematoda	272	Jex et al.	2011
<i>Trichinella spiralis</i>	Pork worm	Nematoda	64	Mitrevva et al.	2011
<i>Anolis carolinensis</i>	North American green anole lizard	Reptile	1,780	Alfoldi et al.	2011
<i>Gasterosteus aculeatus</i>	Freshwater stickleback	Fish	463	Jones et al.	2012
<i>Heliconius melpomene</i>	Postman butterfly	Insect	269	Dasmahapatra et al.	2012
<i>Gorilla gorilla</i>	Western lowland gorilla	Mammal	3,040	Scally et al.	2012
<i>Pan paniscus</i>	Bonobo	Mammal	2,700	Prüfer et al.	2012
<i>Bos grunniens</i>	Domestic yaks	Mammal	2,657	Qiu et al.	2012
<i>Schistosoma haematobium</i>	Blood fluke	Trematoda	385	Young et al.	2012

karyotes, including humans. A decade ago, the publication of the *D. melanogaster* complete genome sequence, with 120 Mb genome and 13,600 genes, spurred a strong uptick in research of fruit fly genomics (Adams et al., 2000). Afterwards, a second genome of *Drosophila* species, *D. pseudoobscura*, was published in 2005 (Richards et al., 2005) and additional ten genomes (*sechellia*, *simulans*, *yakuba*, *erecta*, *ananassae*, *persimilis*, *willistoni*, *mojavensis*, *virilis*, and *grimshawi*) of *Drosophila* species were dispersed a further two years later (Richards et al., 2005; Clark et al., 2007).

Comparative analyses of 12 genomes in a phylogenetic framework illustrated how rates and patterns of sequence divergence across taxa can illuminate evolutionary processes on a genomic scale (Clark et al., 2007). The genomes of *Drosophila* are remarkably conserved across species, with very similar features such as overall genome size, number of genes, distribution of transposable element classes, and patterns of codon usage. Even so, many variable characteristics were also disclosed across these taxa (Clark et al., 2007). Firstly, there are abundant genomic structural changes and rearrangements, particularly exemplified by several different rearrangements of *Hox* gene clusters. Secondly, gene family size is highly variable, with almost half of all gene families changed in size on at least one lineage and with a noticeable fraction of rapid and lineage-specific expansions and contractions. Moreover, evolutionary analyses revealed numerous non-neutral changes in functional genomic elements, such as protein-coding genes, non-coding RNA genes, and cis-regulatory regions (Clark et al., 2007).

3.3 Sea urchins

The genome of the long-lived sea urchin *Strongylocentrotus purpuratus*, a model for developmental and systems biology, was published in 2006, providing an evolutionary outgroup for chordates and yielding substantial insights into the evolution of deuterostomes (Sodergren et al., 2006). Genomic analyses facilitated a great number of key discoveries regarding the biology of the sea urchin. Notably, the 23,300 genes in the sea urchin represent nearly all vertebrate gene families. Interestingly, the sea urchin possesses an extensive defense, and orthologs of many human disease-associated genes, etc..

3.4 Frogs

The only published genome sequence in frogs is western clawed frog *Xenopus tropicalis*, and its detailed comparative genomic analyses shed light on the study of embryonic development (Hellsten et al., 2010). Like-

wise, this genome encodes more than 20,000 protein-coding genes, including orthologs of at least 1700 human disease genes, representing 79% of those currently identified in humans. Conservation of the vertebrate immune system is highlighted by genomic comparisons between mammals and *Xenopus* (Robert et al., 2009; Hellsten et al., 2010). Notably, unique antimicrobial peptides possessed by frogs seem to play an important role in skin secretions, a feature absent in birds, reptiles, and mammals. Accordingly, *X. tropicalis*, with more tractable genetics when compared to the widely used model African clawed frog *X. laevis*, became an alternative important model to the latter for studying the vertebrate development.

3.5 Mice, rats and opossums

As ones of the most important model laboratory mammals, the house mouse *Mus musculus* (Chinwalla et al., 2002) and brown Norway rat *Rattus norvegicus* (Rat Genome Sequencing Project Consortium 2004) were among the first mammals to have their genomes deciphered, after humans. Comparative genomic analyses of these species have illuminated some key issues, including the conservation of large-scale synteny across most of the genomes, the proportion of the genomes under selection, number and evolution of protein-coding genes, and the expansion of gene families related to specific adaptations.

Subsequently, the gray short-tailed opossum *Monodelphis domestica* was sequenced, providing a unique perspective on the organization and evolution of mammalian genomes (Mikkelsen et al., 2007). Distinctive features of the opossum chromosomes provide support for recent theories about genome evolution and function, including a strong influence of biased gene conversion on nucleotide sequence composition, and a relationship between chromosomal characteristics and X chromosome inactivation. Comparison of opossum and eutherian genomes (e.g. of mouse, dog, and human) also revealed a sharp difference in evolutionary innovation between protein-coding and non-coding functional elements.

To mine the genomic basis of the unusual physiology and longevity of the naked mole rat *Heterocephalus glaber*, its genome was deciphered in 2011 (Kim et al., 2011). Comparative sequence analyses with the above described three relatives, as well as other mammals (dog, monkey, and human), revealed unique genome features and molecular adaptations consistent with cancer resistance, poikilothermy, hairlessness and insensitivity to low oxygen, altered visual function, circadian rhythms

and taste sensing.

3.6 Macaques

The most frequently used non-human primates in biomedical research are from the genus *Macaca*, which is closely related to humans by virtue of sharing a last common ancestor ~25 mya. To date, three important macaques whose genome sequences have deciphered include the traditional research model Indian rhesus macaque *Macaca mulatta mulatta* (Rhesus Macaque Genome Sequencing and Analysis Consortium 2007), the Chinese rhesus macaque *Macaca mulatta lasiota* and the cynomolgus/crab-eating macaque *Macaca fascicularis* (Yan et al., 2011).

Initial comparative genome analyses of primates including the Indian rhesus macaque, chimpanzees and humans (Rhesus Macaque Genome Sequencing and Analysis Consortium 2007) revealed the structure of ancestral primate genomes and identified evidences for positive selection and lineage-specific expansions and contractions of gene families. Afterwards, comparisons across the three macaques (Yan et al., 2011) revealed that each macaque maintained abundant genetic heterogeneity, including millions of single-nucleotide substitutions and many insertions, deletions and gross chromosomal rearrangements. Genetic divergence patterns suggest that the cynomolgus macaque genome has been shaped by introgression after hybridization with the Chinese rhesus macaque.

Macaque genes display a high degree of sequence similarity with human disease gene orthologs and drug targets. However, Yan and his colleagues also identify several putatively dysfunctional genetic differences between the three macaques (Yan et al., 2011), which may explain some of the functional differences between each group as has been previously observed in clinical studies.

3.7 Apes

Given the many challenges posed to conducting researches on great apes, it is hard to say if they are model organisms, but they are the closest primate relatives of humans and serve as a good model to understand our own species. All great apes and humans belong to the biological superfamily *Hominoidea* that split ~14 mya (*Pongo*), ~7 mya (*Gorilla*), and 3-5 mya (*Pan & Homo*).

To date, four ape species' genome sequences have been deciphered: the Sumatran orangutan *Pongo abelii*, the western lowland gorilla *Gorilla gorilla gorilla*, the common chimpanzee *Pan troglodytes*, and the bonobo *Pan paniscus*. These sequences provided an unprecedented opportunity to teach us about ourselves, both in

terms of the similarities and differences between humans and apes. Sequence comparison between chimpanzee and human genomes generated a largely complete catalogue of the accumulated genetic differences since divergence from our common ancestor (Mikkelsen et al., 2005).

Scientists have likewise revealed that the orangutan genome has many unique features as compared to other primates (Locke et al., 2011), namely that the structural evolution of the orangutan genome has proceeded much more slowly than among other great apes. Besides, a comparison of protein-coding genes revealed approximately 500 genes showing accelerated evolution on each of the gorilla, human and chimpanzee lineages, as well as evidence for parallel acceleration, particularly of the genes involved in hearing (Sally et al., 2012). Furthermore, scientists found that patterns of evolution in human and chimpanzee protein-coding genes are highly correlated and dominated by the fixation of neutral and slightly deleterious alleles.

Notably, the most recently published great ape genome sequence was that of the bonobo (Prüfer et al., 2012). Evolutionary comparative analyses revealed that more than 3% of the human genome is more closely related to either the bonobo or the chimpanzee genome than the latter are to each other. Analysis of these regions will allow various aspects of these two ape species' ancestry to be reconstructed.

4 Other Animals

4.1 Insects

Insects are among the most diverse group of creatures on the planet, representing over half of all known living organisms and over 90% of differing metazoan life forms on earth. To date, we have sequenced the whole genome of at least 32 insect species, meaning that insects represent over 37% of all currently sequenced animals (Fig. 1). Sequences obtained include many important insects such as mosquitoes, ants silkworms, butterflies, wasps, bees, beetles, the human body louse, and the like.

Mosquitoes, another different kind of dipteran insects from classical model fruit flies, are usually considered a nuisance because most of them not only bite living vertebrates for blood-feeding but consequently transmit some of the most harmful human and livestock diseases such as malaria, yellow fever and dengue fever, etc.. Currently, the whole genome for three blood-feeding mosquitoes has been sequenced. The sequence of *Anopheles gambiae*, a principal vector of malaria,

showed marked bimodal density distribution and prominent expansions in specific families of proteins that are likely involved in cell adhesion and immunity (Holt et al., 2002). Surprisingly, the draft genome sequence of another mosquito *Aedes aegypti*, the primary vector for yellow fever and dengue fever, show a genome size ~ 5 times greater than that of *Anopheles gambiae* (Nene et al., 2007a). Nearly 50% of the *A. aegypti* genome consists of transposable elements, which contribute to an approximately 4–6 factor increase in average gene length and in sizes of intergenic regions relative to *A. gambiae*. The third mosquito to have its genome sequenced was the southern house mosquito *Culex quinquefasciatus* (Arensburger et al., 2010); interestingly, its repertoire of 18,883 protein-coding genes is even 22% larger than that of *A. aegypti* and 52% larger than that of *A. gambiae* with multiple gene-family expansions, including olfactory and gustatory receptors, salivary gland genes, and genes associated with xenobiotic detoxification.

As eusocial insects, organized societies of ants include short-lived worker castes that display specialized behaviors alongside long-lived queens dedicated to reproduction. Currently, at least 7 ant species are available with whole genome sequences. Initial genomic comparisons of two socially divergent ant species (*Camponotus floridanus* and *Harpegnathos saltator*) and the gene expression analyses in different castes (Bonasio et al., 2010) provided clues as to the underlying molecular differences across taxa: up-regulation of telomerase and sirtuin deacetylases related to in longer-lived reproductive queens, caste-specific expression of microRNAs and SMYD histone methyltransferases, and differential regulation of genes implicated in neuronal function and chemical communication. Genomes of the Argentine ant *Linepithema humile* and the red harvester ant *Pogonomyrmex barbatus* dispersed in a year later (Smith et al., 2011b) showed distinctive features including remarkable gene family expansions, an abundance of cytochrome P450 genes, and complete CpG DNA methylation toolkits. Comparative genomic analyses of the fire ant *Solenopsis invicta*, another major pest, reveal an ancestral vitellogenin gene that first underwent a duplication and then was followed by independent duplications of the daughter vitellogenin genes and subfunctionalization with queen- and worker-specific expression (Wurm et al., 2011). Genomic comparison of two leafcutter ants (*Atta cephalotes* and *Acromyrmex echinator*) revealed insights into their key adaptations to advanced social life and fungus farming during obligate

symbiotic lifestyle (Nygaard et al., 2011; Suen et al., 2011).

The first reported butterfly genome was for the migratory monarch butterfly *Danaus plexippus*, which yielded insights into the genetic and molecular basis of long-distance migration (Zhan et al., 2011), involved with circadian clockwork, regulation of the juvenile hormone biosynthetic pathway, oriented flight behavior, differential expression of microRNA between summer and migratory (winter) forms and monarch-specific expansions of chemoreceptors. Additionally, genome sequencing analyses of another butterfly *Heliconius melpomene*, together with two co-mimics (*H. timareta* and *H. elevatus*) via resequencing, reveals a promiscuous exchange of mimicry adaptations among species of the rapidly radiating genus of neotropical butterflies, especially at two genomic regions that control mimicry patterns (Dasmahapatra et al., 2012).

4.2 Fishes

The term "fish" most precisely describes any non-tetrapod craniate that consists of all gill-bearing aquatic craniate animals lacking limbs with digits. To date, at least 5 fish draft genome sequences have been published.

The first of these sequenced genomes is from the fish commonly known as "torafugu" *Fugu rubripes*. Of the 365-Mb pufferfish genome, repeats account for $< 1/3$ while gene sequences occupying $\sim 1/3$ of the genome (Aparicio et al., 2002). Although 75% of predicted human proteins have a strong match to *Fugu*, approximately 25% of predicted human proteins have highly diverged from or have no pufferfish homologs, highlighting the great extent of protein evolution in the 450 myr since teleosts and mammals diverged.

The second fish genome comes from the teleost fish *Tetraodon nigroviridis* (Jaillon et al., 2004), which is a freshwater pufferfish with the smallest known vertebrate genome, at 340Mb. Comparison with other vertebrates and an urochordate indicates that fish proteins diverged markedly faster than their mammalian homologues. Analysis of the *Tetraodon* and human genomes shows that whole-genome duplication occurred in the teleost fish lineage, subsequent to its divergence from tetrapods.

Genomic analyses of a small egg-laying freshwater teleost, medaka *Oryzias latipes*, show a strict genetic separation of 4 myr between two inbred strains derived from two regional populations, and furthermore suggests that differential selective pressures acted on specific gene categories (Kasahara et al., 2007). Compari-

sons with the human, pufferfish *Tetraodon*, zebrafish and medaka genomes revealed that eight major inter-chromosomal rearrangements took place in the remarkably short period beginning approximately 50 myr after the whole-genome duplication event in the teleost ancestor and continuing onward thereafter.

Genome sequence analyses of the Atlantic cod *Gadus morhua*, a cold-adapted teleost which sustains long-standing commercial fisheries and incipient aquaculture, shows evidence for complex thermal adaptations in its haemoglobin gene cluster as well as an unusual immune architecture compared with other sequenced vertebrates (Star et al., 2011). The major histocompatibility complex (MHC) II is a conserved feature of the adaptive immune system of jawed vertebrates, but the Atlantic cod has lost the genes for MHC II, CD4 and invariant chain (Ii) that are essential for the function of this pathway. Results also showed a highly expanded number of MHC I genes and a unique composition of its toll-like receptor families.

Stickleback fishes have colonized and adapted to thousands of streams and lakes formed since the last ice age. With a high-quality reference genome assembly of a threespine stickleback *Gasterosteus aculeatus* and sequencing data of an additional twenty individuals from a global set of marine and freshwater populations (Jones et al., 2012), researchers identified that the reuse of globally shared standing genetic variation, including coding changes, regulatory changes, and chromosomal inversions, plays an important role in repeated evolution of distinct marine and freshwater sticklebacks, and in the maintenance of divergent ecotypes during the early stages of reproductive isolation.

4.3 Birds

With around 10,000 living species, birds (class Aves) are the most speciose class of tetrapod vertebrates. Now, at least three important species of birds have had their sequenced genomes with publications. The first sequenced genome of birds is from chicken, i.e., the red jungle-fowl *Gallus gallus*, a modern descendant of the dinosaurs, and its genome sequence provides new perspectives on vertebrate evolution (Hillier et al., 2004). For instances, the evolutionary dynamics of protein domains and orthologous groups in coding regions illustrates lineage-specific processes leading to both birds and mammals. Likewise, distinctive properties of avian micro-chromosomes together with the inferred patterns of conserved synteny provide additional insights into vertebrate chromosome architecture.

The second sequenced avian whole genome sequence

is from the zebra finch *Taeniopygia guttata*, a songbird and an important model organism in several fields with unique relevance to human neuroscience. The overall genomic structures of zebra finch are similar to those of chicken's, but differences exist in such characteristics as intra-chromosomal rearrangements, lineage-specific gene family expansions, the number of long-terminal-repeat-based retrotransposons, and mechanisms of sex chromosome dosage compensation (Warren et al., 2010). Song behavior of the songbird engages with gene regulatory networks in its brain, via altering the expression of long ncRNAs, microRNAs, transcription factors and their targets.

The domestic turkey *Meleagris gallopavo*, was the third bird to have its genome sequenced (Dalloul et al., 2010). Comparative genomic analyses among these birds as well as between avian and mammalian support the characteristic stability of avian genomes and identifies genes unique to the avian lineage. Clear differences are seen in number and variety of genes of the avian immune system where expansions and novel genes are less frequent than examples of gene loss.

4.4 Mammals

Apart from models mentioned above, sequenced mammals also include five important domestic animals (dogs, cats, horses, cattle, and yak), four popular and phylogenetically important animals (giant pandas, kangaroos, platypus, and the Tasmanian devil) and an extinct Woolly Mammoth.

A draft genome sequence of the domestic dog *Canis familiaris* was reported together with a dense map of SNPs across various breeds with great phenotypic diversity in morphological, physiological and behavioral traits (Lindblad-Toh et al., 2005). Sequence comparison with the primate and rodent lineages shed light on the structure and evolution of genomes and genes. Notably, the majority of the most highly conserved non-coding sequences in mammalian genomes are clustered near a small subset of genes with important roles in development. Analysis of SNPs reveals long-range haplotypes across the entire dog genome, and defines the nature of genetic diversity within and across breeds.

The genome for an inbred Abyssinian domestic cat *Felis catus* (Pontius et al., 2007) illustrated historic balancing of translocation and inversion incidences in distinct mammalian lineages, suggested by annotated features, including repetitive elements, endogenous retroviral sequences, nuclear mitochondrial sequences, micro-RNAs, and evolutionary breakpoints.

The genome sequence of the low-altitude domesti-

cated cow *Bos taurus*, an important livestock for milk and meat production, opened a window to ruminant biology and evolution (Elsik et al., 2009). Results of the sequencing indicate that cattle-specific evolutionary breakpoint regions in chromosomes have a higher density of segmental duplications, enrichment of repetitive elements, and species-specific variations in genes associated with lactation and immune responsiveness. Genomic comparisons between cattle and the domestic yak *Bos grunniens*, which lives at a high altitude on the Qinghai-Tibetan Plateau and near the adjacent regions, provided informative insights into their adaptation to high altitude (Qiu et al., 2012). These comparative genomic analyses also allow identifying expansion of yak gene families related to sensory perception and energy metabolism, as well as an enrichment of protein domains involved in sensing the extracellular environment and hypoxic stress. Positively selected and rapidly evolving genes in the yak lineage are also found to be significantly enriched in functional categories and pathways related to hypoxia and nutrition metabolism.

In 2009, the draft genome sequence of the horse *Equus caballus*, one of the earliest domesticated species and one that has played a vital role in human exploration of novel territories and territorial expansion, was completed (Wade et al., 2009). Comparative analyses showed that chromosomes appear to have undergone few historical rearrangements: 53% of equine chromosomes show conserved synteny to a single human chromosome. Equine chromosome 11 was shown to have an evolutionary new centromere devoid of centromeric satellite DNA, suggesting that centromeric function may have arisen before satellite repeat accumulation.

The genome sequence of the tammar wallaby *Macropus eugenii*, a member of the kangaroo family (Renfree et al., 2011), meanwhile, provides novel insight into the evolution of mammalian reproduction, development and genome evolution. To illustrate the point, scientists identified innovations in reproductive and lactational genes, rapid evolution of germ cell genes, and incomplete, locus-specific X inactivation.

Genomic analyses of the platypus *Ornithorhynchus anatinus*, a monotreme, an order that exhibits a fascinating combination of reptilian and mammalian characters, revealed several unique signatures of animal evolution. The analysis of the first monotreme genome highlighted how reptile and platypus venom proteins have been co-opted independently from the same gene families, while milk protein genes are conserved despite

platypuses laying eggs, and the immune gene family expansions are directly related to platypus biology (Warren et al., 2008).

A draft genome sequence of the much adored but equally endangered giant panda *Ailuropoda melanoleura* was published with detailed genomic analyses (Li et al., 2009). This special creature has several unusual biological and behavioral traits: a restrictive diet of bamboo, a very low fecundity, and a phylogenetic position in continuing controversy. Comparisons with other mammals revealed that the panda genome has not greatly diverged from either dogs or humans, but that there has been considerable divergence in the repetitive regions, most of which seems a result from recent transposable-element activity. The assessment of panda genes had some considerable insights that may bode well for conservation efforts, as genomic analysis suggested that its unique bamboo diet may be more dependent on its gut microbiome than its own genetic composition.

4.5 Other animals

Alongside the species mentioned above, other 15 animals with publications on their whole genome projects (Table 1) listed as following: three blood flukes (*Schistosoma mansoni*, *S. japonicum*, and *S. haematobium*), a North American green anole lizard *Anolis carolinensis*, sea anemone *Nematostella vectensis*, coral *Acropora digitifera*, red red spider mite *Tetranychus urticae*, vase tunicate *Ciona intestinalis*, water flea *Daphnia pulex*, sponge *Amphimedon queenslandica*, hydra *Hydra magnipapillata*, trypanosome *Trypanosoma cruzi*, lancelet *Branchiostoma floridae*, trichoplax *Trichoplax adhaerens*. Comparative genomics analyses have also been conducted more or less for these animals.

5 Conclusions

This review marks one of many preludes to what will certainly be an explosion in the field of animal genomics over the coming years, though we may still overlook some excellent studies of evolutionary genomics. As we march down the path of deciphering more and more animals' whole-genome sequences, genomic analyses has and will continue to shed new light on the evolutionarily and economically important genomics components across the kingdom *Animalia* as well as refinement of the current phylogenetic tree obtained by fossils and other methods. These progresses will also make more and more animals potential model for particular interests in life sciences, just as the jumping ant

H. saltator and the naked mole rat *H. glaber* for longevity models or *X. tropicalis* for an alternative amphibian model for vertebrate embryonic development. Thus, the insightful evolutionary genomics of animals, together with their physiology, morphology, and behavior, will unambiguously begin the process of ushering basic and applied biology into a new era, one marked by a more truly systematic knowledge of organisms surrounding us.

Acknowledgments This work was supported by an National Natural Science Foundation of China key grant (No. 30930056) and a Chinese Academy of Sciences key grant (KSCX2-YW-N-060) to W.W..

References

- Abad P, Gouzy J, Aury JM, Castagnone-Sereno P, Danchin EG et al., 2008. Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nat. Biotechnol.* 26: 909–915.
- Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD et al., 2000. The genome sequence of *Drosophila melanogaster*. *Science* 287: 2185–2195.
- Alfoldi J, Di Palma F, Grabherr M, Williams C, Kong L et al., 2011. The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* 477: 587–591.
- Aparicio S, Chapman J, Stupka E, Putnam N, Chia J et al., 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* 297: 1301–1310.
- Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P et al., 2010. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. *Science* 330: 86–88.
- Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP et al., 2009. The genome of the blood fluke *Schistosoma mansoni*. *Nature* 460: 352–358.
- Bonasio R, Zhang G, Ye C, Mutti NS, Fang X et al., 2010. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science* 329: 1068–1071.
- Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T et al., 2010. The dynamic genome of *Hydra*. *Nature* 464: 592–596.
- Chinwalla AT, Cook LL, Delehaunty KD, Fewell GA, Fulton LA et al., 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520–562.
- Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B et al., 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450: 203–218.
- Colbourne JK, Pfrender ME, Gilbert D, Thomas WK, Tucker A et al., 2011. The ecoresponsive genome of *Daphnia pulex*. *Science* 331: 555–561.
- C. elegans* Sequencing Consortium, 1998. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* 282: 2012–2018.
- Dalloul RA, Long JA, Zimin AV, Aslam L, Beal K et al., 2010. Multi-platform next-generation sequencing of the domestic turkey *Meleagris gallopavo*: Genome assembly and analysis. *PLoS Biol* 8: e1000475.
- Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A et al., 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487: 94–98.
- Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B et al., 2002. The draft genome of *Ciona intestinalis*: Insights into chordate and vertebrate origins. *Science* 298: 2157–2167.
- Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K et al., 2008. The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nat Genet* 40: 1193–1198.
- El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G et al., 2005. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309: 409–415.
- Elsik CG, Tellam RL, Worley KC, 2009. The genome sequence of taurine cattle: A window to ruminant biology and evolution. *Science* 324: 522–528.
- Ghedini E, Wang S, Spiro D, Caler E, Zhao Q et al., 2007. Draft genome of the filarial nematode parasite *Brugia malayi*. *Science* 317: 1756–1760.
- Grbic M, Van Leeuwen T, Clark RM, Rombauts S, Rouze P et al., 2011. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. *Nature* 479: 487–492.
- Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J et al., 2010. The genome of the Western clawed frog *Xenopus tropicalis*. *Science* 328: 633–636.
- Hillier LDW, Miller W, Birney E, Warren W, Hardison RC et al., 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432: 695–716.
- Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R et al., 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 298: 129–149.
- Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N et al., 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431: 946–957.
- Jex AR, Liu S, Li B, Young ND, Hall RS et al., 2011. *Ascaris suum* draft genome. *Nature* 479: 529–533.
- Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E et al., 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484: 55–61.
- Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W et al., 2007. The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447: 714–719.
- Kim EB, Fang X, Fushan AA, Huang Z, Lobanov AV et al., 2011. Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature* 479: 223–227.
- Kirkness EF, Haas BJ, Sun W, Braig HR, Perotti MA et al., 2010. Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc Natl Acad Sci U S A* 107: 12168–12173.
- Li R, Fan W, Tian G, Zhu H, He L et al., 2009. The sequence and de novo assembly of the giant panda genome. *Nature* 463:

- 311–317.
- Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB et al., 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438: 803–819.
- Locke DP, Hillier LDW, Warren WC, Worley KC, Nazareth LV et al., 2011. Comparative and demographic analysis of orang-utan genomes. *Nature* 469: 529–533.
- Mikkelsen TS, Hillier LDW, Eichler EE, Zody MC, Jaffe DB et al., 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437: 69–87.
- Mikkelsen TS, Wakefield MJ, Aken B, Amemiya CT, Chang JL et al., 2007. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature* 447: 167–177.
- Miller W, Drautz DI, Ratan A, Pusey B, Qi J et al., 2008. Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* 456: 387–390.
- Miller W, Hayes VM, Ratan A, Petersen DC, Wittekindt NE et al., 2011. Genetic diversity and population structure of the endangered marsupial *Sarcophilus harrisii* (Tasmanian devil). *Proceedings of the National Academy of Sciences* 108: 12348–12353.
- Mitreva M, Jasmer DP, Zarlenga DS, Wang Z, Abubucker S et al., 2011. The draft genome of the parasitic nematode *Trichinella spiralis*. *Nat Genet* 43: 228–235.
- Nene V, Wortman JR, Lawson D, Haas B, Kodira C et al., 2007b. Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* 316: 1718–1723.
- Nygaard S, Zhang G, Schiott M, Li C, Wurm Y et al., 2011. The genome of the leaf-cutting ant *Acromyrmex echinator* suggests key adaptations to advanced social life and fungus farming. *Genome Res* 21: 1339–1348.
- Opperman CH, Bird DM, Williamson VM, Rokhsar DS, Burke M et al., 2008. Sequence and genetic map of *Meloidogyne hapla*: A compact nematode genome for plant parasitism. *Proc. Natl. Acad. Sci. USA* 105: 14802–14807.
- Pontius JU, Mullikin JC, Smith DR, Team AS, Lindblad-Toh K et al., 2007. Initial sequence and comparative analysis of the cat genome. *Genome research* 17: 1675–1689.
- Prüfer K, Munch K, Hellmann I, Akagi K, Miller JR et al., 2012. The bonobo genome compared with the chimpanzee and human genomes. *Nature* 486: 527–531.
- Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U et al., 2008. The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453: 1064–1071.
- Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J et al., 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317: 86–94.
- Qiu Q, Zhang G, Ma T, Qian W, Wang J et al., 2012. The yak genome and adaptation to life at high altitude. *Nat. Genet.* 44: 946–949.
- Rat Genome Sequencing Project Consortium, 2004. Genome sequence of the brown Norway rat yields insights into mammalian evolution. *Nature* 428: 493–521.
- Renfree MB, Papenfuss AT, Deakin JE, Lindsay J, Heider T et al., 2011. Genome sequence of an Australian kangaroo *Macropus eugenii* provides insight into the evolution of mammalian reproduction and development. *Genome Biology* 12: R81.
- Rhesus Macaque Genome Sequencing and Analysis Consortium, 2007. Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 316: 222–234.
- Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R et al., 2008. The genome of the model beetle and pest *Tribolium castaneum*. *Nature* 452: 949–955.
- Richards S, Liu Y, Bettencourt BR, Hradecky P, Letovsky S et al., 2005. Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. *Genome research* 15: 1–18.
- Robert J, Ohta Y, 2009. Comparative and developmental study of the immune system in *Xenopus*. *Developmental Dynamics* 238: 1249–1270.
- Scally A, Dutheil JY, Hillier LDW, Jordan GE, Goodhead I et al., 2012. Insights into hominid evolution from the gorilla genome sequence. *Nature* 483: 169–175.
- Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K et al., 2011. Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* 476: 320–323.
- Smith CD, Zimin A, Holt C, Abouheif E, Benton R et al., 2011a. Draft genome of the globally widespread and invasive Argentine ant *Linepithema humile*. *Proc Natl Acad Sci U S A* 108: 5673–5678.
- Smith CR, Smith CD, Robertson HM, Helmkampf M, Zimin A et al., 2011b. Draft genome of the red harvester ant *Pogonomyrmex barbatus*. *Proc. Natl. Acad. Sci. USA* 108: 5667–5672.
- Sodergren E, Weinstock GM, Davidson EH, Cameron RA, Gibbs RA et al., 2006. The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* 314: 941–952.
- Srivastava M, Begovic E, Chapman J, Putnam NH, Hellsten U et al., 2008. The *Trichoplax* genome and the nature of placozoans. *Nature* 454: 955–960.
- Srivastava M, Simakov O, Chapman J, Fahey B, Gauthier ME et al., 2010. The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature* 466: 720–726.
- Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrøm M et al., 2011. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477: 207–210.
- Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR et al., 2003. The genome sequence of *Caenorhabditis briggsae*: A platform for comparative genomics. *PLoS Biol* 1: E45.
- Suen G, Teiling C, Li L, Holt C, Abouheif E et al., 2011. The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle. *PLoS Genet* 7: e1002007.
- The Honeybee Genome Sequencing Consortium, 2006. Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443: 931–949.
- The International Aphid Genomics Consortium, 2010. Genome sequence of the pea aphid *Acyrtosiphon pisum*. *PLoS Biol* 8: e1000313.
- The Schistosoma japonicum Genome Sequencing and Functional Analysis Consortium, 2009. The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature* 460: 345–351.

- Venkatesh B, Kirkness EF, Loh YH, Halpern AL, Lee AP et al., 2007. Survey sequencing and comparative analysis of the elephant shark *Callorhynchus milii* genome. *PLoS Biol* 5: e101.
- Wade C, Giulotto E, Sigurdsson S, Zoli M, Gnerre S et al., 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* 326: 865–867.
- Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LDW et al., 2010. The genome of a songbird. *Nature* 464: 757–762.
- Warren WC, Hillier LDW, Graves JAM, Birney E, Ponting CP et al., 2008. Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 453: 175–183.
- Werren JH, Richards S, Desjardins CA, Niehuis O, Gadau J et al., 2010. Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science* 327: 343–348.
- Wurm Y, Wang J, Riba-Grognuz O, Corona M, Nygaard S et al., 2011. The genome of the fire ant *Solenopsis invicta*. *Proc. Natl. Acad. Sci. USA* 108: 5679–5684.
- Xia Q, Zhou Z, Lu C, Cheng D, Dai F et al., 2004. A draft sequence for the genome of the domesticated silkworm *Bombyx mori*. *Science* 306: 1937–1940.
- Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z et al., 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.* 29: 735–741.
- Yan G, Zhang G, Fang X, Zhang Y, Li C et al., 2011. Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. *Nat. Biotechnol.* 29: 1019–1023.
- Young ND, Jex AR, Li B, Liu S, Yang L et al., 2012. Whole-genome sequence of *Schistosoma haematobium*. *Nat. Genet.* 44: 221–225.
- Zhan S, Merlin C, Boore JL, Reppert SM, 2011. The monarch butterfly genome yields insights into long-distance migration. *Cell* 147: 1171–1185.