



Review in Advance first posted online on November 20, 2013. (Changes may still occur before final publication online and in print.)

# Domestication Genomics: Evidence from Animals

Guo-Dong Wang, Hai-Bing Xie, Min-Sheng Peng, David Irwin, and Ya-Ping Zhang

State Key Laboratory of Genetic Resources and Evolution and Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China; email: zhangyp@mail.kiz.ac.cn

Annu. Rev. Anim. Biosci. 2014. 2:24.1–24.20

The *Annual Review of Animal Biosciences* is online at [animal.annualreviews.org](http://animal.annualreviews.org)

This article's doi:  
10.1146/annurev-animal-022513-114129

Copyright © 2014 by Annual Reviews.  
All rights reserved

## Keywords

domestication genomics, origin, demographic history, artificial selection

## Abstract

Animal domestication has far-reaching significance for human society. The sequenced genomes of domesticated animals provide critical resources for understanding the genetic basis of domestication. Various genomic analyses have shed a new light on the mechanism of artificial selection and have allowed the mapping of genes involved in important domestication traits. Here, we summarize the published genomes of domesticated animals that have been generated over the past decade, as well as their origins, from a phylogenomic point of view. This review provides a general description of the genomic features encountered under a two-stage domestication process. We also introduce recent findings for domestication traits based on results from genome-wide association studies and selective-sweep scans for artificially selected genomic regions. Particular attention is paid to issues relating to the costs of domestication and the convergent evolution of genes between domesticated animals and humans.

## INTRODUCTION

The domestication of wild animals by human beings was a great revolution. Domesticated animals dramatically changed human life, allowing a shift from a hunter-gatherer society to a farming society, and were a prerequisite for the rise of human civilization. Since the agricultural revolution of the early Neolithic period, approximately 12,000–14,000 years ago (1), humans have attempted to tame wild animals and breed them through a process of artificial selection, cultivating their roles as food sources (e.g., meat and milk), commodity producers (e.g., wool and silk), or workers (e.g., transportation and protection). Some animals, such as the cat, rat, and house sparrow, became closely associated with humans during the Neolithic age. However, there does not appear to have been a conscious effort to domesticate these species; rather, their domestication occurred via natural selection (2). More than 40 animal species were domesticated in several geographic areas, known as the centers of domestication (3). Some species underwent a single domestication event within a restricted geographic area (e.g., dog and donkey), whereas others underwent multiple independent domestications in different regions (e.g., pig and chicken) (4). Long-standing animal husbandry and controlled breeding have made the behavior and morphological characteristics of domestic animals quite different from their wild ancestors and have greatly shaped the genetic diversity among breed populations.

The abundance of phenotypic variation seen in domesticated animals provided the seed that led to Darwinian evolutionary biology (5). Darwin documented variation under domestication (6), and his observations raised a series of fundamental questions: Where, when, and how did these domesticated animals originate? What is the genetic basis of the domestication process? What is the influence of artificial selection, and what are the differences between natural and artificial selection? Today, sequencing and assembly of genomes is a remarkable strategy for advancing our understanding of domestication through the development of genome technologies, such as the whole-genome shotgun (WGS) approach (7), next-generation massively parallel sequencing (NGS) (8, 9), and even third-generation single-molecule sequencing technologies (10). Genome-sequencing projects of domesticated animals not only offer very powerful resources for answering the questions raised by Darwin but also offer opportunities to explore the genetic basis of economic traits. In this review, we attempt to summarize the current knowledge on genome projects of domesticated animals and discuss the origins, demographic history, and artificial selection of these species. Furthermore, we conclude with a preview of future directions for this field.

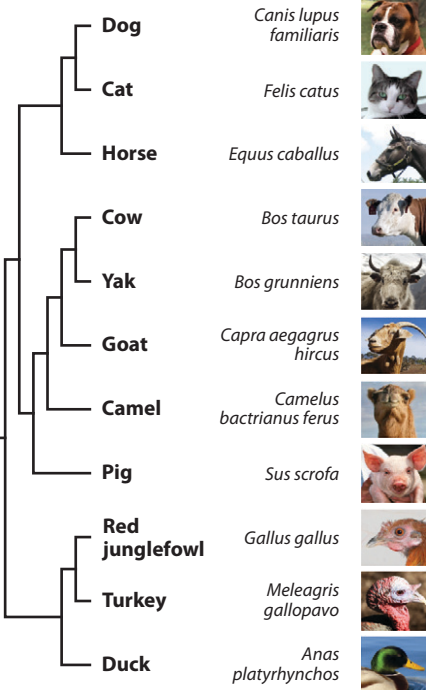
## GENOMES OF DOMESTICATED ANIMALS












Several technologies were developed to allow de novo assembly of genomes during the human genome project (11, 12). WGS sequencing, together with the construction of physical maps, was the most effective approach. By 2009, the genomes of four domesticated animals (dog, cat, taurine cattle, and horse) and one genome of a wild species (the red junglefowl, the major ancestor of the chicken) were sequenced and assembled based on this approach (Figure 1).

In 2004, the genome of the red junglefowl (*Gallus gallus*), a major ancestor of the chicken (*Gallus gallus domesticus*), was sequenced. The reference genome was assembled with ~6.6-fold coverage of WGS reads generated by Sanger sequencing technology. Sequencing generated 7,486 contigs with a N50 size of 36 kb, and 37 supercontigs (also defined as scaffolds) were generated with a N50 size of 7.07 Mb. Of the 1.05 Gb of assembled genome sequence, 933 Mb were localized to specific chromosomes facilitated by a physical map for the domesticated chicken (13). It is worth noting that this was the first genome sequence of a wild relative for a domesticated animal (14). As human's best friend, the domesticated dog (*Canis lupus familiaris*) was one of the first domestic animals (15). The genome of a female from the boxer breed was sequenced and assembled using

Wang et al.





			Assembly procedure	Sequence coverage	Genome size (Gb)	N50 contig (kb)	N50 scaffold (Mb)	Ref. #
<b>Dog</b>	<i>Canis lupus familiaris</i>		Sanger + physical map	7.5×	2.4	180.0	45.0	7
<b>Cat</b>	<i>Felis catus</i>		Sanger + physical map	1.9×	2.4	20.6	4.7	26
<b>Horse</b>	<i>Equus caballus</i>		Sanger + physical map	6.8×	2.4	112.0	42.0	24
<b>Cow</b>	<i>Bos taurus</i>		Sanger + physical map	9.5×	2.6	103.8	N/A	20
<b>Yak</b>	<i>Bos grunniens</i>		NGS only	65.0×	2.7	20.4	1.4	27
<b>Goat</b>	<i>Capra aegagrus hircus</i>		NGS + restriction map	65.6×	2.7	18.7	3.1	8
<b>Camel</b>	<i>Camelus bactrianus ferus</i>		NGS only	76.0×	2.4	85.3	2.0	28
<b>Pig</b>	<i>Sus scrofa</i>		NGS + physical map	41.0×	2.6	80.7	0.6	30
<b>Red junglefowl</b>	<i>Gallus gallus</i>		Sanger + physical map	6.6×	1.1	36.0	7.1	13
<b>Turkey</b>	<i>Meleagris gallopavo</i>		NGS only	30.0×	1.1	12.6	1.5	9
<b>Duck</b>	<i>Anas platyrhynchos</i>		NGS + physical map	64.0×	1.1	26.0	1.2	32

**Figure 1**

Summary of the published genomes of 11 domesticated animals. The phylogenetic relationship of the domesticated species is shown on the left, with a summary of the sequencing and assembly method, genome size and characteristics, and reference provided on the right. Abbreviation: NGS, next-generation sequencing.

the WGS approach in 2005 (7). A total of 31.5 million sequence reads provided 7.5-fold coverage of the genome, covering 99% of the euchromatic genome. The genome of an inbred Abyssinian domesticated cat (*Felis catus*), another important pet, was produced by the International Cat Genome Sequencing Consortium (16). The latest assembly (*Felis\_catus-6.2*) was updated in September 2011 and is composed of 19 chromosomes with a N50 scaffold length of 4.7 Mb and a N50 contig length of 20.6 kb (<http://www.ensembl.org>). Mullikin et al. (17) provided a large collection of single-nucleotide polymorphisms (SNPs) that mapped across the cat genome through the light-whole-genome-sequence strategy, allowing the discovery of over three million SNPs. These genome and SNP data provide an unparalleled opportunity to explore the genetic basis of the variation in morphological and behavioral traits, and even diseases, in a domesticated animal (18, 19).

Taurine cattle, *Bos taurus*, are the most common type of large, domesticated ungulate. Cattle were domesticated as livestock, for meat; as dairy animals, for milk and other dairy products; and as draft animals (11). In 2009, the domesticated cow became the first livestock animal to have a genome assembly (20). A 2.6-Gb genome was assembled using approximately 9.5-fold coverage sequencing data. The domestication of the horse, *Equus caballus*, enhanced transportation and warfare capabilities, which played an important role in human exploration of novel territories

(21–23). The genome of a thoroughbred mare, which is considered to be the epitome of equine athleticism, was sequenced using a WGS approach. A high-quality draft genome sequence of the horse (6.8-fold coverage) was released with a 112-kb N50 contig size and a 46-Mb N50 scaffold size (24).

WGS, together with a physical map, is an effective approach to assemble a genome sequence; however, generating WGS reads by Sanger sequencing technology is expensive. At present, NGS technologies provide an alternative but effective way for whole-genome sequencing because of their improved sequencing throughput and reduced costs (25, 26). The NGS strategy was proven to be successful with the sequencing of the genome of the giant panda (26), and it has since been used to sequence the genomes of three domesticated animals: turkey (*Meleagris gallopavo*) (9), yak (*Bos grunniens*) (27), and camel (*Camelus bactrianus ferus*) (28).

However, genome assemblies using NGS technology cannot locate contigs and scaffolds onto chromosomes, thus limiting the completeness and accuracy of the resulting genome assemblies (29). De novo assembly of genomes from only short-read data is a huge bioinformatic challenge; thus, the construction of a physical map is still extremely helpful. For instance, the assembly of the pig (*Sus scrofa domestica*) genome, which was constructed with Illumina short-read data produced by the Swine Genome Sequencing Consortium (30), was supplemented by physical maps generated from bacterial artificial chromosome clones (31). The pig genome (Sscrofa10.2) comprises 2.60 Gb of sequence assigned to 18 autosomes and two sex chromosomes. Similarly, the duck (*Anas platyrhynchos*) genome was also assembled based on this combined approach (32).

Construction of physical maps is, unfortunately, a laborious task. A new whole-genome mapping approach has been developed that generates a highly contiguous assembly for a large genome without the aid of traditional physical maps (8). First, the Illumina sequencing platform was used to generate 191.5 Gb of reads for assembling contigs and scaffolds. Second, the high-throughput whole-genome mapping instrument produced single-molecule restriction maps for generating superscaffolds using a hybrid assembly approach. Finally, the superscaffolds were anchored to the 29 autosomes and X chromosome based on chromosomal colinearity. This procedure was successfully used to assemble the goat (*Capra hircus*) genome (8).

The genomes of additional domestic animals will be sequenced, such as the sheep (*Ovis aries*) genome that is being produced by the International Sheep Genomics Consortium and is now online at Ensembl (<http://pre.ensembl.org/>). With reductions in the costs of sequencing and assembly genomes, additional domesticated species and their wild relatives will be sequenced, providing important genetic resources. The pairing of genomes from domesticated animals and their wild ancestors is very powerful for understanding the demographic history of domesticated animals, providing new insight into the fine mapping of traits associated with domestication, as well as shedding new light on the mechanisms and consequences of artificial selection at the genomic level.

## ORIGINS OF DOMESTICATED ANIMALS

The origin of domestic animals is an old but pivotal question (6). As a landmark for the agricultural revolution, zooarchaeological approaches have been included in studies of human evolution (33). During the past twenty years, molecular genetic approaches have been applied to investigate this issue (34, 35). Researchers can infer the likely progenitor of the domesticated species through phylogenetic analyses of extant domesticated species and their wild relatives (36, 37). Incorporation of geographic information of the samples allows further biogeographic or phylogeographic analyses (38), which can discern the center(s) of domestication. Samples from these centers are expected to have higher haplotypic and nucleotide diversity and harbor more ancestral variation or unique lineages (36, 37). Under the hypothesis of a molecular clock (39), the timescales

Wang et al.



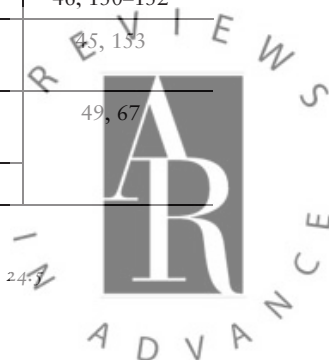
for domestication events can be estimated based on known calibration points or the mutation rates of genetic markers (36, 37).

Compared with nuclear markers, mitochondrial DNA (mtDNA) has some properties (e.g., lack of recombination, high mutation rate, multiple copies) that make it useful for tracing the origin of domesticated animals (36, 40). Since 2000, a large number of studies have used sequences of mtDNA fragments [i.e., control region (D-loop) and/or *cytochrome b*] to study various domesticated animals (3), including species that have worldwide distribution, such as pigs (41, 42), horses (23, 43), cattle (44), goats (45), sheep (46), dogs (47), donkeys (48), chickens (49), and cats (50), as well as those with more limited distributions, such as zebu cattle (51), water buffalo (52), yaks (53, 54), and alpacas and llamas (55). Results from mtDNA studies provide a perspective, at least from the maternal side, as to the likely progenitors and candidate lineages involved in domestication. The inferred domestication centers or events are listed in **Table 1**.

With progress in the efficiency of DNA sequencing, researchers began to adopt a population phylogenomic approach using complete mitochondrial genomes rather than just fragments. Compared with just the control region, the sequencing of many complete mtDNA sequences allowed

**Table 1** Origin of eight common domesticated animals inferred by mitochondrial DNA

Clade	Domesticated species		Wild ancestor		Location	References
	Name	Latin	Name	Latin		
Mammals	Dog	<i>Canis lupus familiaris</i>	Gray wolf	<i>Canis lupus</i>	Southern East Asia	47, 62, 69
	Cat	<i>Felis silvestris catus</i>	African wildcat	<i>Felis silvestris lybica</i>	West Asia	50
	Donkey	<i>Equus asinus</i>	African wild ass	<i>Equus africanus</i>	Northeast Africa	48
	Pig	<i>Sus scrofa domestica</i>	Wild boar	<i>Sus scrofa</i>	East Asia	41, 57, 149
					South Asia	
					Southeast Asia	
					West Asia	
					Europe	
	Cattle	<i>Bos taurus</i>	Aurochs	<i>Bos primigenius</i>	West Asia	44, 58–61
					North Africa	
Europe?						
Horse	<i>Equus caballus</i>	Wild horse	<i>Equus ferus</i>	Eurasian steppes	23, 43, 58, 64	
Sheep	<i>Ovis aries</i>	Asian mouflon	<i>Ovis orientalis</i>	West Asia	46, 150–152	
Goat	<i>Capra aegagrus hircus</i>	Bezoar	<i>Capra aegagrus</i>	West Asia	45, 153	
Birds	Chicken	<i>Gallus gallus domesticus</i>	Red junglefowl	<i>Gallus gallus</i>	Southern East Asia	49, 67
					South Asia	



clarification of mtDNA phylogeny/genealogy of domesticated species at the highest levels of resolution, a strategy that has proved successful in molecular anthropological research (56). This approach was first carried out for pigs in 2007 (57) and then became widely adopted for taurine cattle (58–61), dogs (62), horses (63, 64), sheep (65), yaks (66), and chickens (67). With their high resolution and increased sampling sizes, these studies yield more details about the origin of domestication, including the time of origin, number of founders, and approximate geographical region. These approaches have also revealed some previously unknown, cryptic local demographic events or sporadic introgression from wild progenitors (57–61, 67).

An analog of mtDNA, the Y chromosome (strictly, only the male-specific portion of the Y chromosome) can reflect the parental contribution to the gene pool of mammalian domesticates (40, 68). Integrating information from the independently inherited mtDNA and Y chromosome genetic markers can provide greater details about the origins of domesticated species. For instance, early work based on 654-mtDNA control-region sequences suggested that dogs had a single origin in the vast region of Asia east of the Urals (47). Subsequent analyses of 169 mitochondrial genomes, together with 1,543 control region sequences, further refined the origin of dogs to a location south of the Yangtze River in East Asia approximately 5,400–16,300 years ago (62). By analyzing the sequence variation of the Y chromosome in 151 dogs, Ding et al. (69) corroborated the conclusion that the area south of the Yangtze River was the principal, and possibly sole, region of dog domestication, but they also revealed that subsequent dog-wolf hybridization events contributed modestly to the dog gene pool. Here, both the Y chromosome and the mtDNA analyses presented strikingly similar pictures for dog domestication, with the place and time coinciding approximately with the origin of rice agriculture, suggesting that dogs may have originated among sedentary hunter-gatherers or early farmers as an important cultural trait (62, 69). Unlike the pattern observed in dogs, inconsistent results derived from Y chromosome and mtDNA data may indicate sex-biased breeding histories (40). In contrast to a high diversity observed with mtDNA (23, 63, 64), near-complete monomorphism of the Y chromosome is found in extant domestic horses (70, 71). These results suggest that modern horses trace back to an extremely limited number of male lineages despite having multiple maternal origins, implying a strong sex-biased domestication history (70). Unfortunately, compared with mtDNA, systematic investigations of the Y chromosome in domesticated species are still limited (68), in part because the sequence of the Y chromosome is often unknown, as most reference genome sequences for domesticated species are from females (i.e., they lack a Y chromosome).

Both mtDNA and Y chromosome markers are useful for tracing the origin and subsequent dispersal of domesticated animals, and this information is often valuable for other genetic studies. Nevertheless, traditional Sanger sequencing makes massive analyses of mitochondrial genomes and Y chromosomes at the population level cumbersome and expensive. Targeted resequencing of specific portions of mtDNA and Y chromosomes using NGS (72–74) provided an opportunity to address this issue. The initial studies using this approach involved the resequencing of mitochondrial genomes (64) and Y chromosomes (75) in horses. Resequencing of approximately 186 kb of the horse Y chromosome identified three Y-chromosomal haplotypes (75) and resolved an important piece of the genetic puzzle of horse breeds. Additional data from mtDNA and Y-chromosome markers generated using NGS platforms are expected to enhance our genetic documentation of the genealogy of various domesticated animals in the near future.

There are limits to the knowledge that can be gained from mtDNA and Y-chromosome markers. The nonrecombining mtDNA and Y chromosome can be treated as two independent genetic loci in the genome (36), but they reflect just a small part in whole-genomic history (76). In addition, the effective population size for haploid mtDNA and Y chromosomes corresponds to a quarter of the size of the diploid autosomes. mtDNA and Y-chromosome markers are sensitive to

Wang et al.





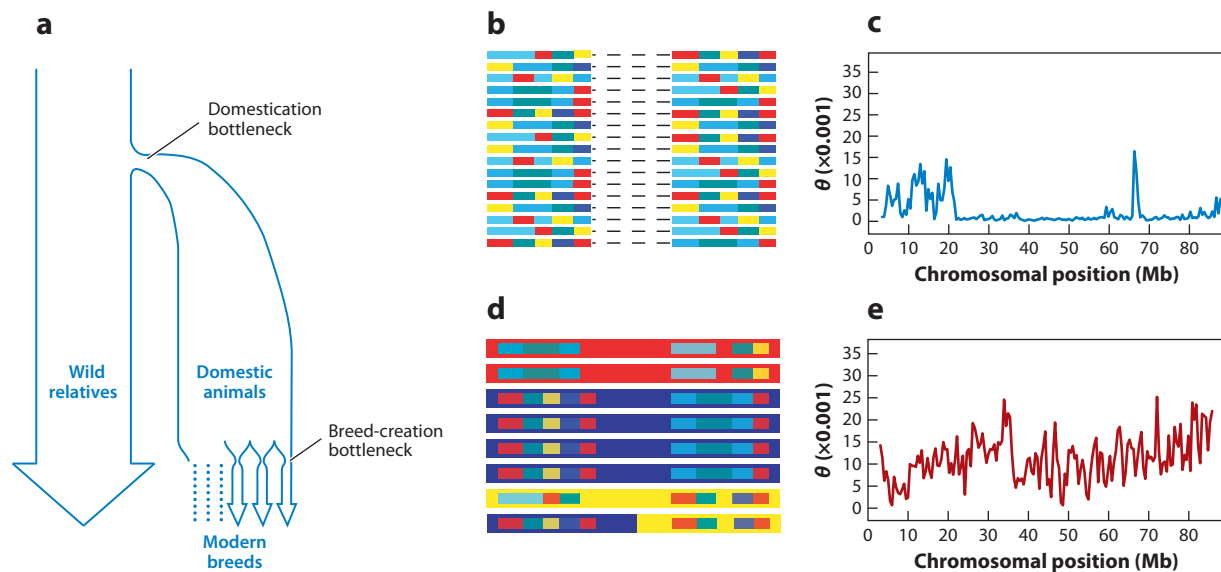
genetic drift (77); moreover, natural selection acting on the two markers cannot be excluded (76, 77). As a result of these limitations, it is not surprising that nuclear autosomal markers revealed different scenarios for the origin of some domesticated species. For instance, work on the *BCDO2* (*beta-carotene dioxygenase 2*) gene indicated that *yellow skin* originated not from the red junglefowl, as suggested by mtDNA (Table 1), but most likely from the gray junglefowl (*Gallus sonneratii*) (78). With advances in genotyping and resequencing technologies, analyses of autosomal genome-wide markers can be employed. For instance, whole-genome resequencing of dogs revealed that the split between wolves and Chinese indigenous dogs occurred 32,000 years ago (79), much earlier than estimated from mtDNA (47, 62). vonHoldt et al. (80) showed that dog breeds share a higher proportion of multilocus haplotypes unique to gray wolves from the Middle East, suggesting that this location was the dominant source of the genome diversity seen in dogs. Incorporation of data from mtDNA, the Y chromosome, and autosomes of extensive samples, even ancient DNAs, should allow us to depict a more complete picture for the origin of domesticated dogs. Domestication can mirror the movement of humans during the Neolithic era, as domesticated animals were transported by or with humans. One paradigm is the commensal models used to discern the peopling of the Pacific (81). Analyses of mtDNA variation in chickens (82, 83), dogs (84), and pigs (85, 86) revealed that Southeast Asia was an important source for the Pacific domesticates. This suggested a protracted period of genetic exchange between the exotic breeds carried by the Austronesians [presumably from Taiwan (87)] and the indigenous animals of Southeast Asia, implying that the Austronesian immigrants probably sojourned in Southeast Asia before spreading to the Pacific, which would support the slow-boat model for the populating of the Pacific (88). Similarly, work with the *MC1R* (*melanocortin receptor1*) gene revealed selection for the black color in China, which could be associated with the ancient Chinese sacrificial culture, providing an interesting example of the molecular consequences of artificial selection in livestock and its cultural impact (89).

## GENOMIC FEATURES UNDER THE DOMESTICATION PROCESS

Genomic patterns often show that the evolution of domesticated animals proceeds through a two-stage process (Figure 2a). Generally, two evolutionary stages can be partitioned in many domesticated animals: (a) an ancient domestication event followed by (b) a more recent breeding process. Ancient domestication events correspond to the domestication of the animal from its wild ancestors to a domesticated species; many such events occurred approximately 12,000–14,000 years ago, during the agricultural revolution of the early Neolithic, along with the domestication of major crops (3). Compared with the more ancient domestication events, modern breeding formation has been largely restricted to the past 300 years. Breed formation started with selective breeding of animals in response to diverse demands in the eighteenth century (90). For example, in the domesticated pig, the selective goals were directed at higher growth rates and production of leaner meat in response to changes in human caloric intake in modern societies (91). The selective goals for dog domestication have been more diverse than for any other domesticated animals, as they have been bred not only for desired physical characteristics, such as body size, skull shape, coat color, texture, length, and thickness, and even tail shape and size (90, 92, 93), but also for behavioral patterns, with breeds specialized for herding, guarding, agility, speed, and companionship (94, 95).

Domesticated animals underwent an ancient bottleneck when they were domesticated from their wild relatives. Population genomic analyses showed that the strength of these ancient bottlenecks was relatively mild (7, 24). For instance, the domesticated dog experienced a relatively mild bottleneck during which the effective population size was reduced to 16% of the ancestral





**Figure 2**

Demographic history and genomic feature under a two-stage domestication history. (a) Domesticated animals underwent an ancient domestication bottleneck and more recent breed-creation bottlenecks (154). The strength of the ancient bottleneck was relatively mild, leading to (b) short-range linkage disequilibrium (LD) (7) and (c) relatively low nuclear diversity (96). The breed-creation bottlenecks were of high strength and resulted in (d) long-range breed-specific haplotypes (7) and (e) extremely low nuclear diversity (96).

population size (96). As a result of these ancient bottlenecks, domesticated species had only short-range linkage disequilibrium (LD) before the breed-formation processes started (see Figure 2b). Haploblocks in the ancestral dog population were shorter than those in modern humans (~10 kb versus ~20 kb in human). Domesticated pigs from China also have similar LD patterns (~10 kb), reflecting the weak effect of the ancient bottleneck from the early domestication stage (97). LD in horses across breeds is intermediate in length between that seen in dogs and in humans (24). Nuclear diversity in domesticated animals before the breeding process started was also mildly reduced compared with their wild relatives owing to the mild ancient bottleneck (Figure 2c). For instance, the reduction of diversity in the ancient dog population was approximately 80% of that of gray wolves (96), and mitochondrial evidence revealed that domesticated pigs show lower levels of nucleotide diversity than do wild boars (57).

### GENOMIC FEATURE UNDER THE BREED-FORMATION PROCESS

The recent and intensive inbreeding that occurred during breed formation has introduced a strong bottleneck effect in modern breeds (for dogs, see Figure 2a). The short time span of selection has resulted in breed-specific, long-range haplotypes, which have not yet been substantially broken down by genetic recombination. Consequently, strong LD is expected in breeds that have been exposed to modern breeding.

For example, LD extends for ~2 Mb in modern dog breeds, which is approximately 10–50-times longer than that observed in most human populations (98). Long haplotypes are very likely the result of inbreeding rather than the target of selection. As shown in Figure 2d, long haplotypes in dogs are composed of concatenations of several short ancestral haplotypes, which have early

Wang et al.



origins from the ancient population and are shared by many breeds (7). Domesticated dogs have relatively long LD (~2 Mb) within breeds and much shorter LD (~10 kb) across breeds (7) (**Figure 2b**), which is attributed to genetic isolation during the breeding process. A different situation is found in horses, where LD across breeds is only slightly shorter than that observed within breeds (100–150 kb). A possible explanation for this difference is that a large mare population was required for breed generation owing to the small number of offspring produced per mare (24). Amaral et al. (97) investigated three genomic regions in a wide diversity of domesticated pigs (from both Europe and China) and French wild boars and found that the LD extended (with  $r^2 \geq 0.3$ ) over large genomic haploblocks of up to 400 kb for most European domestic pigs, while being only approximately 10 kb in most Chinese domesticated pigs. The extensive recent breeding history in European pigs potentially explains this difference in LD size. The high intensity of inbreeding that occurred in European domesticated pigs would have led to increased homozygosity (identity-by-descent) and, as a consequence, an elevated level of LD.

In addition to the enrichment of long haplotypes across the whole genome, another feature of recent breed formation is that these specialized breeds should suffer from severe founder effects. A reduction of genetic diversity is expected from the historic bottlenecks that occurred during modern breed formation. An example is shown in **Figure 2c,e**, where a significant reduction in genetic diversity would be expected after intensive breeding. In the swine industry, for example, many globally distributed breeds were created from a few European domesticated pigs in the past 300 years. In contrast, many local pig breeds have been well retained in China (99, 100). After investigating 39 independent genomic markers in a wide range of pig populations, Megens et al. (101) showed that most European pig breeds exhibit much lower genetic diversity than do the local Chinese pig breeds. This difference has two possible explanations: Reduced genetic diversity (*a*) could have been an inherited trait owing to a scarcity of polymorphisms in the ancestral population or (*b*) could be due to a severe founder effect as a result of the recent breeding process. A smaller effective population size for the ancestral European wild boars has been suggested, owing to a strong population bottleneck during the Last Glacial Maximum (30). However, the reduction is still observed if European pig breeds are combined together and compared with small subsets of the Chinese pig breeds that suffered from domestication bottlenecks owing to their geographically restricted regions (101). Therefore, the data suggest that the severe founder effects generated during modern breeding explain the differences in observed genetic diversity between European and Chinese domesticated pigs.

## GENOME-WIDE ASSOCIATION AND WHOLE-GENOME RESEQUENCING STUDIES TO EXPLORE COMPLEX TRAITS

Domesticated animals offer an opportunity to study the genetic mechanisms involved in complex traits. The recent origin of domesticated animals results in them having a high background level of genetic homogeneity. Through whole-genome sequencing projects, a large number of SNPs have been discovered and can be used in genome-wide association studies (GWAS), and commercial SNP chips are available for common domesticated animals (**Table 2**). Many SNPs have been identified that associate with diverse traits, such as fattiness, androstenone levels, and brown coat color in pigs (102–104); dwarfism in horses (105); abdominal fattiness and body weight in chickens (106, 107); and several diseases in dogs (108–110).

Owing to the limited density of SNPs included in the chips, GWAS studies usually do not specifically identify the responsible gene or causative mutation underlying the phenotypic variation examined. An analysis that narrows down the target region is necessary to refine the association map. For example, Sutter et al. identified a major quantitative trait locus in dogs on chromosome



**Table 2 Available commercial single-nucleotide polymorphism (SNP) chips for genome-wide association studies in domesticated animals**

Species	SNPs	Chip
Cattle	54,609	Illumina BovineSNP50 BeadChip
	777,962	Illumina BovineHD BeadChip
	6,909	Illumina BovineLD BeadChip
	648,855	Axiom Genome-Wide BOS1 Bovine Array
Dog	22,362	Illumina CanineSNP20 BeadChip
	172,115	Illumina CanineHD BeadChip
Sheep	54,241	Illumina OvineSNP50 BeadChip
Chicken	60,000	Illumina ChickenSNP60 BeadChip
	580,961	Axiom Genome-Wide Chicken Genotyping Array
Pig	64,232	Illumina PorcineSNP60 BeadChip
Horse	54,602	Illumina EquineSNP50 BeadChip
Water buffalo	90,000	Axiom Buffalo Genotyping Array

15 within a single breed, the Portuguese water dog, a breed that descended from two major founding kennels that disagreed on an appropriate size standard, resulting in considerable size variation in the founding population (111–113). The gene for insulin-like growth factor 1 (*IGF1*) was identified as an important factor contributing to body-size variation when SNPs and small insertion/deletion polymorphisms in the 15-Mb genomic interval surrounding the quantitative trait locus were examined in small and giant breeds. Further analyses showed that there was a significant signal for intense artificial selection on the *IGF1* locus across 22 breeds of small and giant dogs, suggesting that a single *IGF1* allele is a major determinant of small size in dogs (113).

GWAS provides a fast and standardized procedure to interrogate some complex traits, but several prerequisites are required. First, a relatively large sample size is needed to generate sufficient detection power and have a low false-discovery rate. Second, cosegregating SNPs that will be genotyped on the chips should flank the causative mutation(s) underlying the phenotypic variation examined. The second prerequisite, however, does not always hold true. A substantial amount of phenotypic variation is breed specific; thus, the detection power of GWAS will differ greatly among breeds. SNPs exhibit differing detection powers among populations in GWAS studies, especially those from animals that had multiple domestication origins; therefore, customized chips with increased SNP density should be developed. An alternative and enhanced approach to GWAS is to sequence a whole genome from a population of individuals using NGS platforms. A particular advantage of this method is that it allows the discovery of many breed-specific mutations.

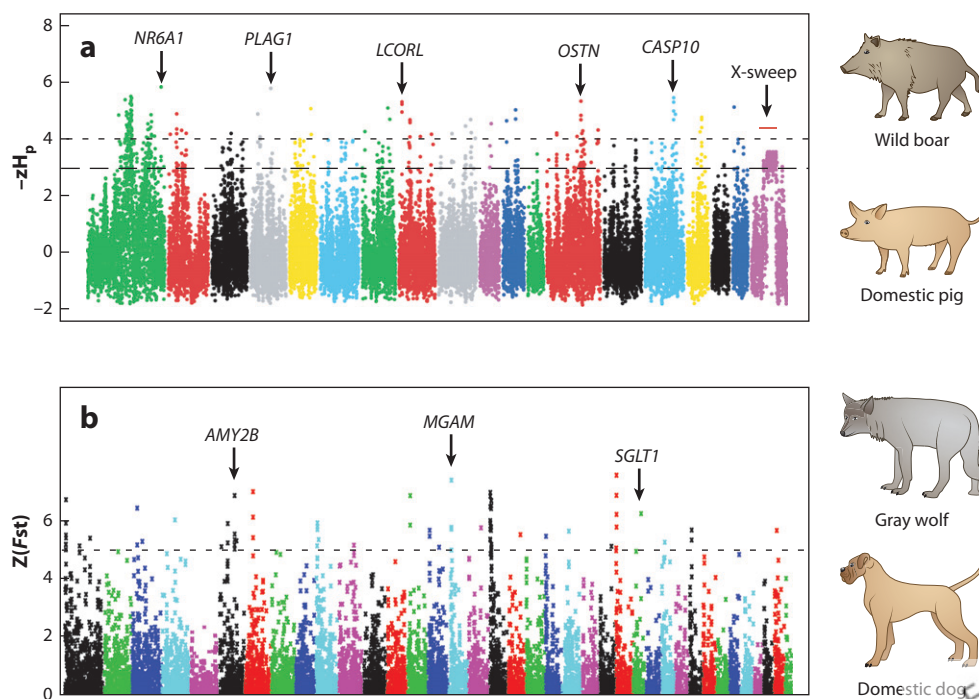
With whole-genome resequencing data, selective sweeps can be detected to identify candidate genetic sequences that underlie complex traits. One signature of a selective sweep is a regional reduction in the nucleotide diversity at the loci under selection. Genome-wide screening of commercial pig breeds showed a cluster of genomic regions that have low nucleotide diversity (114). A frequently used measure of nucleotide diversity is the Watterson's estimator ( $\theta_w$ ) (115). In the Large White, Landrace, and Pietrain breeds of pigs, the *KIT* locus, which controls white coat color pigmentation, shows lower nucleotide diversity ( $\theta_w$  ranging from 0.00021 to 0.00034) compared

Wang et al.



with in wild boars (0.00078) or red Duroc (0.00067) (114). Because it is well documented that the *KIT* polymorphism is associated with the dominant white color in domestic pigs (116, 117), Sutter et al. (113) concluded that artificial selection drove the homogeneity at this loci in the Large White, Landrace, and Pietrain breeds.

An alternative measurement of nucleotide diversity is pooled heterozygosity ( $zH_p$ ) (118).  $zH_p$  can be used in cross-group comparison with whole-genome resequencing data to identify areas showing extremely high levels of homozygosity, genomic sequences that could be genetically responsible for phenotypic characteristics of a chosen group. To investigate loci under selection during chicken domestication, Rubin et al. (118) resequenced pools of genomic DNA covering eight different populations of domestic chickens and red junglefowl on the SOLiD platform. By screening selective sweeps in the pooled sequence data, the authors identified some regions with high levels of fixation, thus potentially under selection, in the domesticated chickens. Combining other types of bioinformatic analyses, as well as genetic evidence, the authors proposed that the *TSHR* locus was likely related to photoperiod control of reproduction, where an absence of strict regulation of seasonal reproduction is seen in the domesticated chicken compared with the red junglefowl (118). Recently, Rubin et al. (119) reported that European domesticated pigs (Large White, Landrace, Duroc, Hampshire; 63 individuals in total) contain three genomic regions with excess homozygosity ( $zH_p < -5.29$ ) associated with selection for an elongated back length and increased number of vertebrae (Figure 3a). Three protein-coding genes, including *NR6A1*,



**Figure 3**

Whole-genome scans for positive selection. (a) A selective sweep analysis between the wild boar and the domestic pig (119). Y-axis values are  $-zH_p$ , genome-wide Z score of heterozygosity plot, and the x-axis shows the position of the windows along each chromosome. (b) A selective sweep analysis between the gray wolf and the domestic dog (121). The positive end of the  $Z(F_{st})$  distribution is plotted along the dog autosomes.

*PLAG1*, and *LCORL*, were mapped to these regions and are likely important factors for these valuable traits.

Recent selection in a population can also lead to high levels of genetic differentiation between populations. *Fst* (F statistic) is frequently used to measure the genetic differentiation between populations (120). With whole-genome resequencing data, a Z transformation of *Fst* can be performed to identify genome-wide outliers that show extremely high levels of differentiation. Erik Axelsson et al. (121) recently identified 36 unique autosomal candidate domestication regions containing 122 genes with low  $zH_p$  and high  $Z(Fst)$  scores in dogs (Figure 3b). These genes were enriched in the pathways involved in starch and fat metabolism. The expression level of two of these candidate genes, *MGAM* and *AMY2B* (as well as the amylase activity of the encoded product of the latter), were shown to be significantly higher in the dog compared with the wolf. These results suggest that changes in these genes are adaptations to eating food with starch during the early stage of domestication.

The selective sweep studies described above are largely based on the DNA-pooling sequencing method, which assumes that all samples contribute equally in the pooling DNA. If the samples are unequal, estimation of allele frequency is affected, though equal genomic DNA contribution is hard to guarantee. An alternative method is to sequence samples from isolated libraries; however, this leads to more complicated models to estimate allele frequency owing to the variation in sequencing depths and coverage. To resolve this question, we recently developed a method to detect selective sweeps from individually sequenced samples from populations. In the new method, we model the process to sample chromosomes and infer the allele frequency distribution at each SNP. Varying sequencing depths are allowed among individual SNPs. The population differentiation is calculated between the two distributions, rather than between two fixed allele frequencies (H.B. Xie & Y.P. Zhang, manuscript in preparation).

## COST OF DOMESTICATION

A common observation of domesticated animals is that they tend to have low reproductive fitness and increased susceptibility to various diseases (122, 123). These disadvantageous legacies are known as “the genetic cost of domestication” (124), and they represent the accumulation of deleterious mutations in the genomes of domesticated species (124). Compared with their free-living wild relatives, domesticated animals do not require a high metabolic efficiency owing to their limited activity as well as their more stable living conditions. As mitochondria are the power factory of the cell, they play key roles in oxidative energy metabolism (125). The relaxation of selective constraint since domestication is hypothesized to be reflected in the coding regions of the mitochondrial genome that encodes energy metabolism genes (126). An excess of (slightly) deleterious mutations, that is, nonsynonymous mutations and mutations in RNA-encoding genes, has been characterized in the mitochondrial genomes of domesticated animals, such as the dog (126–128), pig (127), yak (129), and chicken (127). The low-coverage genomic-sequence data also revealed that a higher frequency of nonsynonymous mutations occurred in dogs compared with wolves (130). In addition to the relaxation of selective constraint (96), the Hill-Robertson interference driven by artificial selection also likely contributes to the observed patterns described above (6). Because domestication events are recent, and the deleterious mutations are newly generated, the accumulation of slightly deleterious mutations in domesticated animals can be explained as a consequence of not having enough time to eliminate them from the population (131–133). In other words, excess numbers of slightly deleterious mutations are prone to found on young branches (i.e., domesticated) in a phylogeny (128, 133, 134). To evaluate the cost of domestication, additional genetic data, improved methods for analyses, and functional experiments will be required.

Wang et al.



## CONVERGENT EVOLUTION

As some of the first to be domesticated, dogs are special animals. Gray wolves, wild relatives of dogs, might have been domesticated first as scavengers that lived and hunted commensally with humans. With successive adaptive changes, humans adopted them during the agricultural revolution (32). Like humans, dogs have adapted to starch-rich diets and turned from being carnivorous to omnivorous. Two independent research teams performing whole-genome resequencing of dogs and wolves have identified several genes with key roles in digestion and metabolism, showing that the large changes in food sources for dogs played a crucial role in the early stages of dog domestication (96, 121). Both studies suggested that the convergent evolution to similar diets caused similar driving forces for positive selection in the genomes of both humans and dogs, resulting in convergent evolution in genes involved in metabolizing this new diet.

As pets, domesticated dogs today are more companions than working animals; thus, they show the greatest level of interspecies bonding with humans (135). Previous research revealed that puppies show skills for tasks for which they must communicate with humans, whereas wolves raised by humans do not. Dogs are even more skillful than great apes (136). They also show ability to learn by mimicking human behaviors. These observations suggest that there has been behavior convergence between humans and dogs. To identify genes that may have been positively selected in this convergence, Wang et al. (96) compared lists of positively selected genes identified in dogs and humans (137) and found that the overlap between these two lists was higher than expected, which revealed that the level of convergence is statistically significant. Interestingly, convergent evolution is apparent not only in genes for digestion and metabolism (121) but also in genes involved in neurological processes (110). Some genes show convergent evolution as being responsible for a wide range of neurological pathogenic conditions, such as aggressive behavior (138), obsessive-compulsive disorder (139), depression, and autism (140). Changes in these neurological genes likely were driven by humans and dogs adapting to their new, similar environments, which had crowded living conditions and required complex intimate interactions both within and between species (141, 142). These results shed new light on the field of convergent evolution and offer opportunities to gain insights into diseases shared by dogs and humans, such as cancer, diabetes, and epilepsy.

## PROSPECTIVE

The genomes of domesticated animals are becoming powerful resources for understanding issues in domestication. With the development of sequencing technologies and assembly tools, additional genomes from various domesticated animals will become available, as will the genomes of their closely related wild relatives and even extinct wild ancestors (143–145). These new genomes will be complemented by genome resequencing at the population level, which should be performed for various indigenous breeds, especially those threatened with replacement by common commercial breeds (146). This will be necessary for the evaluation and conservation of genetic resources. In addition, other genetic approaches, such as transcriptomic (147) and methylomic (148) analyses based on NGS, are required to explore the expression and epigenetic patterns found in wild and domesticated animals to aid in identification of changes associated with domestication. These efforts will not only advance our understanding of the genetic basis of animal domestication but also facilitate improvements through breeding.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.





## ACKNOWLEDGMENTS

We thank Ting-Ting Yin of Kunming Institute of Zoology for drawing figures. This work was supported by the 973 program (2013CB835200, 2013CB835202, 2013CB835204), the National Natural Science Foundation of China (91231108), and the National High Technology Research and Development Program of China (863 Program, 2011AA100304-5). The Youth Innovation Promotion Association, Chinese Academy of Sciences, provided support to M.-S.P.

## LITERATURE CITED

1. Diamond J. 2002. Evolution, consequences and future of plant and animal domestication. *Nature* 418:700–7
2. Driscoll CA, Macdonald DW, O'Brien SJ. 2009. From wild animals to domestic pets, an evolutionary view of domestication. *Proc. Natl. Acad. Sci. USA* 106(Suppl. 1):9971–78
3. Food Agric. Organ. 2007. *The State of the World's Animal Genetic Resources for Food and Agriculture*. Rome: Food Agric. Organ.
4. Diamond J, Bellwood P. 2003. Farmers and their languages: the first expansions. *Science* 300:597–603
5. Darwin C. 1859. *On the Origins of Species by Means of Natural Selection*. London: Murray
6. Darwin C. 1868. *The Variation of Animals and Plants under Domestication*. London: Murray
7. Lindblad-Toh K, Wade CM. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438:803–19
8. Dong Y, Xie M, Jiang Y, Xiao N, Du X, et al. 2012. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nat. Biotechnol.* 31:135–41
9. 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
10. Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, et al. 2012. Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nat. Biotechnol.* 30:693–700
11. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409:860–921
12. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. 2001. The sequence of the human genome. *Science* 291:1304–51
13. Wallis JW, Aerts J, Groenen MA, Crooijmans RP, Layman D, et al. 2004. A physical map of the chicken genome. *Nature* 432:761–64
14. Hillier LW, 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
15. Clutton-Brock J 1995. Origins of the dog: domestication and early history. In *The Domestic Dog: Its Evolution, Behaviour and Interactions with People*, ed. J Serpell, pp. 7–20. Cambridge, UK: Cambridge Univ. Press
16. Pontius JU, Mullikin JC, Smith DR, Lindblad-Toh K, Gnerre S, et al. 2007. Initial sequence and comparative analysis of the cat genome. *Genome Res.* 17:1675–89
17. Mullikin JC, Hansen NF, Shen L, Ebling H, Donahue WF, et al. 2010. Light whole genome sequence for SNP discovery across domestic cat breeds. *BMC Genomics* 11:406
18. Karlsson EK, Lindblad-Toh K. 2008. Leader of the pack: gene mapping in dogs and other model organisms. *Nat. Rev. Genet.* 9:713–25
19. Karlsson EK, Baranowska I, Wade CM, Salmon Hillbertz NHC, Zody MC, et al. 2007. Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat. Genet.* 39:1321–28
20. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, et al. 2009. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* 10:R42
21. Anthony DW, Bogucki P, Comşa E, Gimbutas M, Jovanović B, et al. 1986. The “Kurgan Culture,” Indo-European origins, and the domestication of the horse: a reconsideration [and comments and replies]. *Curr. Anthropol.* 27:291–313

Wang et al.





22. Diamond JM. 1991. The earliest horse men. *Nature* 350:275–76
23. Vila C, Leonard JA, Gotherstrom A, Marklund S, Sandberg K, et al. 2001. Widespread origins of domestic horse lineages. *Science* 291:474–77
24. Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, et al. 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* 326:865–67
25. Wang J, Wang W, Li R, Li Y, Tian G, et al. 2008. The diploid genome sequence of an Asian individual. *Nature* 456:60–65
26. Li R, Fan W, Tian G, Zhu H, He L, et al. 2010. The sequence and de novo assembly of the giant panda genome. *Nature* 463:311–17
27. 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.* 44:946–49
28. Bactrian Camels Genome Seq. Anal. Consort. 2012. Genome sequences of wild and domestic bactrian camels. *Nat. Commun.* 3:1202
29. Lewin HA, Larkin DM, Pontius J, O'Brien SJ. 2009. Every genome sequence needs a good map. *Genome Res.* 19:1925–28
30. Groenen MA, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, et al. 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491:393–98
31. Humphray SJ, Scott CE, Clark R, Marron B, Bender C, et al. 2007. A high utility integrated map of the pig genome. *Genome Biol.* 8:R139
32. Huang Y, Li Y, Burt DW, Chen H, Zhang Y, et al. 2013. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat. Genet.* 45:776–83
33. Vigne JD, Peters J, Helmer D, eds. 2005. *The First Steps of Animal Domestication: New Archaeozoological Techniques*. Oxford: Oxbow Books. 176 pp.
34. Loftus RT, MacHugh DE, Bradley DG, Sharp PM, Cunningham P. 1994. Evidence for two independent domestications of cattle. *Proc. Natl. Acad. Sci. USA* 91:2757–61
35. MacHugh DE, Loftus RT, Bradley DG, Sharp PM, Cunningham P. 1994. Microsatellite DNA variation within and among European cattle breeds. *Proc. Biol. Sci.* 256:25–31
36. Bruford MW, Bradley DG, Luikart G. 2003. DNA markers reveal the complexity of livestock domestication. *Nat. Rev. Genet.* 4:900–10
37. Zeder MA, Emshwiller E, Smith BD, Bradley DG. 2006. Documenting domestication: the intersection of genetics and archaeology. *Trends Genet.* 22:139–55
38. Avise JC. 2009. Phylogeography: retrospect and prospect. *J. Biogeogr.* 36:3–15
39. Kumar S. 2005. Molecular clocks: four decades of evolution. *Nat. Rev. Genet.* 6:654–62
40. MacHugh DE, Bradley DG. 2001. Livestock genetic origins: goats buck the trend. *Proc. Natl. Acad. Sci. USA* 98:5382–84
41. Larson G, Dobney K, Albarella U, Fang M, Matisoo-Smith E, et al. 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science* 307:1618–21
42. Giuffra E, Kijas JM, Amarger V, Carlborg Ö, Jeon JT, Andersson L. 2000. The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* 154:1785–91
43. Jansen T, Forster P, Levine MA, Oelke H, Hurler M, et al. 2002. Mitochondrial DNA and the origins of the domestic horse. *Proc. Natl. Acad. Sci. USA* 99:10905–10
44. Troy CS, MacHugh DE, Bailey JF, Magee DA, Loftus RT, et al. 2001. Genetic evidence for Near-Eastern origins of European cattle. *Nature* 410:1088–91
45. Luikart G, Gielly L, Excoffier L, Vigne JD, Bouvet J, Taberlet P. 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc. Natl. Acad. Sci. USA* 98:5927–32
46. Hiendleder S, Kaup B, Wassmuth R, Janke A. 2002. Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc. Biol. Sci.* 269:893–904
47. Savolainen P, Zhang YP, Luo J, Lundeberg J, Leitner T. 2002. Genetic evidence for an East Asian origin of domestic dogs. *Science* 298:1610–13
48. Beja-Pereira A, England PR, Ferrand N, Jordan S, Bakhiet AO, et al. 2004. African origins of the domestic donkey. *Science* 304:1781

49. Liu YP, Wu GS, Yao YG, Miao YW, Luikart G, et al. 2006. Multiple maternal origins of chickens: out of the Asian jungles. *Mol. Phylogenet. Evol.* 38:12–19
50. Driscoll CA, Menotti-Raymond M, Roca AL, Hupe K, Johnson WE, et al. 2007. The Near Eastern origin of cat domestication. *Science* 317:519–23
51. Chen S, Lin BZ, Baig M, Mitra B, Lopes RJ, et al. 2010. Zebu cattle are an exclusive legacy of the South Asia neolithic. *Mol. Biol. Evol.* 27:1–6
52. Kierstein G, Vallinoto M, Silva A, Schneider MP, Iannuzzi L, Brenig B. 2004. Analysis of mitochondrial D-loop region casts new light on domestic water buffalo (*Bubalus bubalis*) phylogeny. *Mol. Phylogenet. Evol.* 30:308–24
53. Guo S, Savolainen P, Su J, Zhang Q, Qi D, et al. 2006. Origin of mitochondrial DNA diversity of domestic yaks. *BMC Evol. Biol.* 6:73
54. Lai SJ, Chen SY, Liu YP, Yao YG. 2007. Mitochondrial DNA sequence diversity and origin of Chinese domestic yak. *Anim. Genet.* 38:77–80
55. Kadwell M, Fernandez M, Stanley HF, Baldi R, Wheeler JC, et al. 2001. Genetic analysis reveals the wild ancestors of the llama and the alpaca. *Proc. Biol. Sci.* 268:2575–84
56. Torroni A, Achilli A, Macaulay V, Richards M, Bandelt HJ. 2006. Harvesting the fruit of the human mtDNA tree. *Trends Genet.* 22:339–45
57. Wu GS, Yao YG, Qu KX, Ding ZL, Li H, et al. 2007. Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. *Genome Biol.* 8:R245
58. Achilli A, Bonfiglio S, Olivieri A, Malusa A, Pala M, et al. 2009. The multifaceted origin of taurine cattle reflected by the mitochondrial genome. *PLoS ONE* 4:e5753
59. Achilli A, Olivieri A, Pellecchia M, Uboldi C, Colli L, et al. 2008. Mitochondrial genomes of extinct aurochs survive in domestic cattle. *Curr. Biol.* 18:R157–58
60. Bonfiglio S, Achilli A, Olivieri A, Negrini R, Colli L, et al. 2010. The enigmatic origin of bovine mtDNA haplogroup R: Sporadic interbreeding or an independent event of *Bos primigenius* domestication in Italy? *PLoS ONE* 5:e15760
61. Bonfiglio S, Jinja C, De Gaetano A, Achilli A, Olivieri A, et al. 2012. Origin and spread of *Bos taurus*: new clues from mitochondrial genomes belonging to haplogroup T1. *PLoS ONE* 7:e38601
62. Pang JF, Kluetsch C, Zou XJ, Zhang AB, Luo LY, et al. 2009. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Mol. Biol. Evol.* 26:2849–64
63. Achilli A, Olivieri A, Soares P, Lancioni H, Hooshiar Kashani B, et al. 2012. Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proc. Natl. Acad. Sci. USA* 109:2449–54
64. Lippold S, Matzke NJ, Reissmann M, Hofreiter M. 2011. Whole mitochondrial genome sequencing of domestic horses reveals incorporation of extensive wild horse diversity during domestication. *BMC Evol. Biol.* 11:328
65. Meadows JR, Hiendleder S, Kijas JW. 2011. Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. *Heredity* 106:700–6
66. Wang Z, Shen X, Liu B, Su J, Yonezawa T, et al. 2010. Phylogeographical analyses of domestic and wild yaks based on mitochondrial DNA: new data and reappraisal. *J. Biogeogr.* 37:2332–44
67. Miao YW, Peng MS, Wu GS, Ouyang YN, Yang ZY, et al. 2013. Chicken domestication: an updated perspective based on mitochondrial genomes. *Heredity* 110:277–82
68. Greminger MP, Krützen M, Schelling C, Pienkowska-Schelling A, Wandeler P. 2010. The quest for Y-chromosomal markers—methodological strategies for mammalian non-model organisms. *Mol. Ecol. Resour.* 10:409–20
69. Ding ZL, Oskarsson M, Ardalan A, Angleby H, Dahlgren LG, et al. 2011. Origins of domestic dog in Southern East Asia is supported by analysis of Y-chromosome DNA. *Heredity* 108:507–14
70. Lindgren G, Backström N, Swinburne J, Hellborg L, Einarsson A, et al. 2004. Limited number of patrilineal lines in horse domestication. *Nat. Genet.* 36:335–36

Wang et al.



71. Lau AN, Peng L, Goto H, Chemnick L, Ryder OA, Makova KD. 2009. Horse domestication and conservation genetics of Przewalski's horse inferred from sex chromosomal and autosomal sequences. *Mol. Biol. Evol.* 26:199–208
72. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, et al. 2010. Target-enrichment strategies for next-generation sequencing. *Nat. Methods* 7:111–18
73. Gunnarsdóttir ED, Li M, Bauchet M, Finstermeier K, Stoneking M. 2011. High-throughput sequencing of complete human mtDNA genomes from the Philippines. *Genome Res.* 21:1–11
74. Wei W, Ayub Q, Chen Y, McCarthy S, Hou Y, et al. 2013. A calibrated human Y-chromosomal phylogeny based on resequencing. *Genome Res.* 23:388–95
75. Wallner B, Vogl C, Shukla P, Burgstaller JP, Druml T, Brem G. 2013. Identification of genetic variation on the horse Y chromosome and the tracing of male founder lineages in modern breeds. *PLoS ONE* 8:e60015
76. Balloux F. 2010. The worm in the fruit of the mitochondrial DNA tree. *Heredity* 104:419–20
77. Meiklejohn CD, Montooth KL, Rand DM. 2007. Positive and negative selection on the mitochondrial genome. *Trends Genet.* 23:259–63
78. Eriksson J, Larson G, Gunnarsson U, Bed'hom B, Tixier-Boichard M, et al. 2008. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet.* 4:e1000010
79. Wang G-d, Zhai W, Yang H-c, Fan R-x, Cao X, et al. 2013. The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nat. Commun.* 4:1860
80. vonHoldt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, et al. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* 464:898–902
81. Storey AA, Clarke AC, Ladefoged T, Robins J, Matisoo-Smith E. 2013. DNA and Pacific commensal models: applications, construction, limitations, and future prospects. *J. Island Coast. Arch.* 8:37–65
82. Miao YW, Peng MS, Wu GS, Ouyang YN, Yang ZY, et al. 2013. Chicken domestication: an updated perspective based on mitochondrial genomes. *Heredity* 110:277–82
83. Storey AA, Athens JS, Bryant D, Carson M, Emery K, et al. 2012. Investigating the global dispersal of chickens in prehistory using ancient mitochondrial DNA signatures. *PLoS ONE* 7:e39171
84. Oskarsson MC, Klutsch CF, Boonyaparakob U, Wilton A, Tanabe Y, Savolainen P. 2012. Mitochondrial DNA data indicate an introduction through Mainland Southeast Asia for Australian dingoes and Polynesian domestic dogs. *Proc. Biol. Sci.* 279:967–74
85. Larson G, Cucchi T, Fujita M, Matisoo-Smith E, Robins J, et al. 2007. Phylogeny and ancient DNA of *Sus* provides insights into Neolithic expansion in Island Southeast Asia and Oceania. *Proc. Natl. Acad. Sci. USA* 104:4834–39
86. Lum JK, McIntyre JK, Greger DL, Huffman KW, Vilar MG. 2006. Recent Southeast Asian domestication and Lapita dispersal of sacred male pseudohermaphroditic “tuskers” and hairless pigs of Vanuatu. *Proc. Natl. Acad. Sci. USA* 103:17190–95
87. Diamond JM. 2000. Taiwan's gift to the world. *Nature* 403:709–10
88. Oppenheimer SJ, Richards M. 2001. Polynesian origins. Slow boat to Melanesia? *Nature* 410:166–67
89. Li J, Yang H, Li JR, Li HP, Ning T, et al. 2010. Artificial selection of the melanocortin receptor 1 gene in Chinese domestic pigs during domestication. *Heredity* 105:274–81
90. Crowley J, Adelman B. 1998. *The Complete Dog Book: Official Publication of the American Kennel Club*. New York: Howell House
91. Merks JW. 2000. One century of genetic changes in pigs and the future needs. *Proc. BSAS Occas. Meet., Chall. Genet. Change Anim. Prod.*, pp. 8–19. Edinburgh, UK: Br. Soc. Anim. Sci.
92. Wayne RK. 1986. Cranial morphology of domestic and wild canids: the influence of development on morphological change. *Evolution* 40:243–61
93. Wayne RK. 1986. Limb morphology of domestic and wild canids: the influence of development on morphologic change. *J. Morphol.* 187:301–19
94. Scott JP, Fuller JL. 1998. *Genetics and the Social Behavior of the Dog*. Chicago: Univ. Chicago Press
95. Ostrander EA, Galibert F, Patterson DF. 2000. Canine genetics comes of age. *Trends Genet.* 16:117–24
96. Wang GD, Zhai W, Yang HC, Fan RX, Cao X, et al. 2013. The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nat. Commun.* 4:1860



97. Amaral AJ, Megens HJ, Crooijmans RP, Heuven HC, Groenen MA. 2008. Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics* 179:569–79
98. Wayne RK, Ostrander EA. 2007. Lessons learned from the dog genome. *Trends Genet.* 23:557–67
99. Porter V. 1993. *Pigs: A Handbook to the Breeds of the World*. Ithaca, NY: Cornell Univ. Press
100. Zhang Z. 1986. *Pig Breeds in China*. Shanghai, China: Shanghai Sci. Tech. Publ.
101. Megens HJ, Crooijmans RP, San Cristobal M, Hui X, Li N, Groenen MA. 2008. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. *Genet. Sel. Evol.* 40:103–28
102. Duijvesteijn N, Knol EF, Merks JW, Crooijmans RP, Groenen MA, et al. 2010. A genome-wide association study on androstene levels in pigs reveals a cluster of candidate genes on chromosome 6. *BMC Genet.* 11:42
103. Ponsuksili S, Murani E, Brand B, Schwerin M, Wimmers K. 2011. Integrating expression profiling and whole-genome association for dissection of fat traits in a porcine model. *J. Lipid Res.* 52:668–78
104. Ren J, Mao H, Zhang Z, Xiao S, Ding N, Huang L. 2011. A 6-bp deletion in the *TYRP1* gene causes the brown colouration phenotype in Chinese indigenous pigs. *Heredity* 106:862–68
105. Orr N, Back W, Gu J, Leegwater P, Govindarajan P, et al. 2010. Genome-wide SNP association-based localization of a dwarfism gene in Friesian dwarf horses. *Anim. Genet.* 41(Suppl. 2):2–7
106. Abasht B, Lamont SJ. 2007. Genome-wide association analysis reveals cryptic alleles as an important factor in heterosis for fatness in chicken F2 population. *Anim. Genet.* 38:491–98
107. Gu X, Feng C, Ma L, Song C, Wang Y, et al. 2011. Genome-wide association study of body weight in chicken F2 resource population. *PLoS ONE* 6:e21872
108. Awano T, Johnson GS, Wade CM, Katz ML, Johnson GC, et al. 2009. Genome-wide association analysis reveals a *SOD1* mutation in canine degenerative myelopathy that resembles amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. USA* 106:2794–99
109. Wood SH, Ke X, Nuttall T, McEwan N, Ollier WE, Carter SD. 2009. Genome-wide association analysis of canine atopic dermatitis and identification of disease related SNPs. *Immunogenetics* 61:765–72
110. Li Y, vonHoldt BM, Reynolds A, Boyko AR, Wayne RK, et al. 2013. Artificial selection on brain-expressed genes during the domestication of dog. *Mol. Biol. Evol.* In press
111. Chase K, Carrier DR, Adler FR, Jarvik T, Ostrander EA, et al. 2002. Genetic basis for systems of skeletal quantitative traits: principal component analysis of the canid skeleton. *Proc. Natl. Acad. Sci. USA* 99:9930–35
112. Chase K, Carrier DR, Adler FR, Ostrander EA, Lark KG. 2005. Interaction between the X chromosome and an autosome regulates size sexual dimorphism in Portuguese Water Dogs. *Genome Res.* 15:1820–24
113. Sutter NB, Bustamante CD, Chase K, Gray MM, Zhao K, et al. 2007. A single *IGF1* allele is a major determinant of small size in dogs. *Science* 316:112–15
114. Amaral AJ, Ferretti L, Megens HJ, Crooijmans RP, Nie H, et al. 2011. Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. *PLoS ONE* 6:e14782
115. Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* 7:256–76
116. Kijas JM, Wales R, Törnsten A, Chardon P, Moller M, Andersson L. 1998. Melanocortin receptor 1 (*MC1R*) mutations and coat color in pigs. *Genetics* 150:1177–85
117. Johansson Moller M, Chaudhary R, Hellmén E, Høyheim B, Chowdhary B, Andersson L. 1996. Pigs with the dominant white coat color phenotype carry a duplication of the *KIT* gene encoding the mast/stem cell growth factor receptor. *Mamm. Genome* 7:822–30
118. Rubin CJ, Zody MC, Eriksson J, Meadows JR, Sherwood E, et al. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464:587–91
119. Rubin CJ, Megens HJ, Martinez Barrio A, Maqbool K, Sayyab S, et al. 2012. Strong signatures of selection in the domestic pig genome. *Proc. Natl. Acad. Sci. USA* 109:19529–36
120. Weir B, Cockerham C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–70
121. Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, et al. 2013. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* 495:360–64

Wang et al.



122. Drögemüller C, Karlsson EK, Hytönen MK, Perloski M, Dolf G, et al. 2008. A mutation in hairless dogs implicates *FOXI3* in ectodermal development. *Science* 321:1462
123. Mosher DS, Quignon P, Bustamante CD, Sutter NB, Mellersh CS, et al. 2007. A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. *PLoS Genet.* 3:e79
124. Lu J, Tang T, Tang H, Huang J, Shi S, Wu CI. 2006. The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication. *Trends Genet.* 22:126–31
125. Wallace DC. 2011. Bioenergetic origins of complexity and disease. *Cold Spring Harb. Symp. Quant. Biol.* 76:1–16
126. Björnerfeldt S, Webster MT, Vilà C. 2006. Relaxation of selective constraint on dog mitochondrial DNA following domestication. *Genome Res.* 16:990–94
127. Hughes AL. 2013. Accumulation of slightly deleterious mutations in the mitochondrial genome: a hallmark of animal domestication. *Gene* 515:28–33
128. Soares P, Abrantes D, Rito T, Thomson N, Radivojac P, et al. 2013. Evaluating purifying selection in the mitochondrial DNA of various Mammalian species. *PLoS ONE* 8:e58993
129. Wang Z, Yonezawa T, Liu B, Ma T, Shen X, et al. 2011. Domestication relaxed selective constraints on the yak mitochondrial genome. *Mol. Biol. Evol.* 28:1553–56
130. Xia Q, Guo Y, Zhang Z, Li D, Xuan Z, et al. 2009. Complete resequencing of 40 genomes reveals domestication events and genes in silkworm (*Bombyx*). *Science* 326:433–36
131. Ho SY, Lanfear R, Bromham L, Phillips MJ, Soubrier J, et al. 2011. Time-dependent rates of molecular evolution. *Mol. Ecol.* 20:3087–101
132. Ho SY, Larson G. 2006. Molecular clocks: when times are a-changin'. *Trends Genet.* 22:79–83
133. Kivisild T, Shen PD, Wall DP, Do B, Sung R, et al. 2006. The role of selection in the evolution of human mitochondrial genomes. *Genetics* 172:373–87
134. Peterson GI, Masel J. 2009. Quantitative prediction of molecular clock and  $K_a/K_s$  at short timescales. *Mol. Biol. Evol.* 26:2595–603
135. Tacon P, Pardoe C. 2002. Dogs make us human. *Nat. Aust.* 27(4):52–61
136. Hare B, Brown M, Williamson C, Tomasello M. 2002. The domestication of social cognition in dogs. *Science* 298:1634–36
137. Akey JM. 2009. Constructing genomic maps of positive selection in humans: Where do we go from here? *Genome Res.* 19:711–22
138. Patkar AA, Berrettini WH, Hoehe M, Thornton CC, Gotthel E, et al. 2002. Serotonin transporter polymorphisms and measures of impulsivity, aggression, and sensation seeking among African-American cocaine-dependent individuals. *Psychiatry Res.* 110:103–15
139. Di Bella D, Erzegovesi S, Cavallini MC, Bellodi L. 2002. Obsessive-compulsive disorder, 5-HTTLPR polymorphism and treatment response. *Pharmacogenomics J.* 2:176–81
140. Brune CW, Kim SJ, Salt J, Leventhal BL, Lord C, Cook EH Jr. 2006. 5-HTTLPR genotype-specific phenotype in children and adolescents with autism. *Am. J. Psychiatry* 163:2148–56
141. de Kloet ER, Joëls M, Holsboer F. 2005. Stress and the brain: from adaptation to disease. *Nat. Rev. Neurosci.* 6:463–75
142. Popova NK. 2006. From genes to aggressive behavior: the role of serotonergic system. *Bioessays* 28:495–503
143. Edwards CJ, Magee DA, Park SD, McGettigan PA, Lohan AJ, et al. 2010. A complete mitochondrial genome sequence from a Mesolithic wild aurochs (*Bos primigenius*). *PLoS ONE* 5:e9255
144. Lari M, Rizzi E, Mona S, Corti G, Catalano G, et al. 2011. The complete mitochondrial genome of an 11,450-year-old aurochs (*Bos primigenius*) from Central Italy. *BMC Evol. Biol.* 11:32
145. Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, et al. 2013. Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499:74–78
146. Hanotte O, Dessie T, Kemp S. 2010. Time to tap Africa's livestock genomes. *Science* 328:1640–41
147. Albert FW, Somel M, Carneiro M, Aximu-Petri A, Halbwax M, et al. 2012. A comparison of brain gene expression levels in domesticated and wild animals. *PLoS Genet.* 8:e1002962
148. Xiang H, Zhu J, Chen Q, Dai F, Li X, et al. 2010. Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. *Nat. Biotechnol.* 28:516–20





149. Larson G, Liu R, Zhao X, Yuan J, Fuller D, et al. 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proc. Natl. Acad. Sci. USA* 107:7686–91
150. Meadows JR, Cemal I, Karaca O, Gootwine E, Kijas JW. 2007. Five ovine mitochondrial lineages identified from sheep breeds of the near East. *Genetics* 175:1371–79
151. Pedrosa S, Uzun M, Arranz JJ, Gutiérrez-Gil B, San Primitivo F, Bayón Y. 2005. Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. *Proc. Biol. Sci.* 272:2211–17
152. Tapio M, Marzanov N, Ozerov M, Činkulov M, Gonzarenko G, et al. 2006. Sheep mitochondrial DNA variation in European, Caucasian, and Central Asian areas. *Mol. Biol. Evol.* 23:1776–83
153. Naderi S, Rezaei HR, Pompanon F, Blum MG, Negrini R, et al. 2008. The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proc. Natl. Acad. Sci. USA* 105:17659–64
154. Wang GD, Cheng LG, Fan RX, Irwin DM, Tang SS, et al. 2013. Signature of balancing selection at the *MC1R* gene in Kunming dog populations. *PLoS ONE* 8:e55469



Wang et al.