Contrasting Patterns of Genomic Diversity Reveal Accelerated Genetic Drift but Reduced Directional Selection on X-Chromosome in Wild and Domestic Sheep Species

Ze-Hui Chen[1,2,†], Min Zhang[1,3,†], Feng-Hua Lv[1,†], Xue Ren[1], Wen-Rong Li[4], Ming-Jun Liu[4], Kiwoong Nam[5], Michael W. Bruford[6], and Meng-Hua Li[1,*]

[1]CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences (CAS), Beijing, China

[2]College of Life Sciences, University of the Academy of Sciences, Beijing 100049, China

[3]School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, China

[4]Animal Biotechnological Research Center, Xinjiang Academy of Animal Science, Urumqi, China

[5]Diversité, Génomes et Interactions Microorganismes Insectes, Institut National de la Recherche Agronomique, University of Montpellier, Montpellier, France

[6]Organisms and Environment Division, School of Biosciences and Sustainable Places Research Institute, Cardiff University, Wales, United Kingdom

[†]These authors contributed equally to this work.

*Corresponding author: E-mail: menghua.li@ioz.ac.cn.

Accepted: April 19, 2018

Abstract

Analyses of genomic diversity along the X chromosome and of its correlation with autosomal diversity can facilitate understanding of evolutionary forces in shaping sex-linked genomic architecture. Strong selective sweeps and accelerated genetic drift on the X-chromosome have been inferred in primates and other model species, but no such insight has yet been gained in domestic animals compared with their wild relatives. Here, we analyzed X-chromosome variability in a large ovine data set, including a BeadChip array for 943 ewes from the world's sheep populations and 110 whole genomes of wild and domestic sheep. Analyzing whole-genome sequences, we observed a substantially reduced X-to-autosome diversity ratio (~0.6) compared with the value expected under a neutral model (0.75). In particular, one large X-linked segment (43.05–79.25 Mb) was found to show extremely low diversity, most likely due to a high density of coding genes, featuring highly conserved regions. In general, we observed higher nucleotide diversity on the autosomes, but a flat diversity gradient in X-linked segments, as a function of increasing distance from the nearest genes, leading to a decreased X: autosome (X/A) diversity ratio and contrasting to the positive correlation detected in primates and other model animals. Our evidence suggests that accelerated genetic drift but reduced directional selection on X chromosome, as well as sex-biased demographic events, explain low X-chromosome diversity in sheep species. The distinct patterns of X-linked and X/A diversity we observed between Middle Eastern and non-Middle Eastern sheep populations can be explained by multiple migrations, selection, and admixture during the domestic sheep's recent postdomestication demographic expansion, coupled with natural selection for adaptation to new environments. In addition, we identify important novel genes involved in abnormal behavioral phenotypes, metabolism, and immunity, under selection on the sheep X-chromosome.

Key words: X-chromosome evolution, *Ovis*, genetic drift, selective sweeps, X/A diversity, whole genomes.

Introduction

The X chromosome and autosomes differ in several aspects of their molecular evolution and population genetics including their effective population sizes (*Ne*), mutation rates, susceptibility to genetic drift, patterns of selection and demography (Schaffner 2004; Vicoso and Charlesworth 2006; Hedrick 2007; Keinan and Reich 2010), all of which differentially affect the genetic diversity of the X-chromosome compared with the autosomes (Begun and Whitley 2000; Casto et al. 2010; Heyer and Segurel 2010) and the X-to-autosome (X/A)

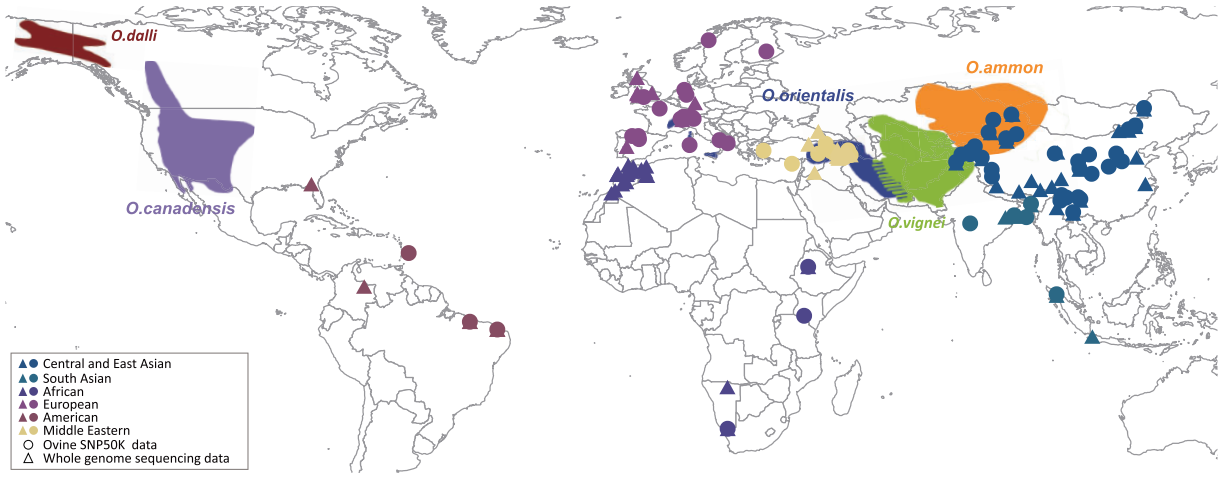

Fig. 1.—Geographic origins of the samples and distribution of the six wild *Ovis* species (*O. orientalis*, *O. ammon*, *O. musimon*, *O. vignei*, *O. dalli*, and *O. canadensis*). Geographic distributions of the six wild *Ovis* species were adapted from figure 1 in Rezaei et al. (2010).

diversity ratio (Ellegren 2009). As a result, deviations from neutral expectation of 0.75 (X)/1.00 (A) (Charlesworth et al. 1987) due to variable diversity ratios across the chromosomes and in specific genomic regions are expected due to the combined effect of different evolutionary forces (Gottipati et al. 2011; Arbiza et al. 2014). Therefore, comparing patterns of X-chromosome and autosomal diversity can provide insight into the relative contributions of these forces in shaping genome evolution.

Several recent studies have compared genomic diversity on the X-chromosome with the autosomes, including comparative analysis of nucleotide diversity as a function of distance from the near genes in humans (Keinan et al. 2009; Hammer et al. 2010; Gottipati et al. 2011; Arbiza et al. 2014) and great apes (Nam et al. 2015, 2017; Narang and Wilson Sayres 2016). More generally, a variety of evolutionary forces have also been invoked to explain patterns of genetic variability on the X chromosome versus autosomes in model species through empirical data analyses (Andolfatto 2001; Lu and Wu 2005; Singh and Petrov 2007; Cox et al. 2010; Corl and Ellegren 2012; VanBuren et al. 2016) and statistical modeling (Betancourt et al. 2004; Ellegren 2009; Heyer and Segurel 2010; Keinan and Reich 2010; Charlesworth 2012; Evans et al. 2014; Veeramah et al. 2014; Lasne et al. 2017).

Unlike most species studied to date, domestic animals have experienced strong and long-term directional selection during the processes of domestication and breeding and they therefore represent an ideal model to explore the evolution of the X chromosome in comparison both with their wild relatives and among divergently selected domestic populations. Sheep (*Ovis aries*) were domesticated in the Fertile Crescent ~ 8000–11000 years ago (Ryder 1983), with diverse native breeds having been developed subsequently across the world with a concomitant demographic expansion, and through adaptation to a diverse range of local environments and varied production systems (Lv et al. 2014). Importantly, their wild relatives (*O. orientalis*, *O. ammon*, *O. musimon*, *O. vignei*, *O. dalli*, *O. canadensis*, and *O. nivicola*) remain extant (Rezaei et al. 2010), making them an ideal model to disentangle the effects of multiple underlying forces on X chromosome evolution.

Here, we analyzed X-chromosomal genetic variability across species in the genus *Ovis*, comprising a large data set for domestic sheep and their wild relatives. The data consist of genome-wide BeadChip SNP genotypes of 943 females from 68 domestic populations (fig. 1 and supplementary table S1, Supplementary Material online) and 110 whole genomes of wild and domestic sheep from around the world (fig. 1 and supplementary table S2, Supplementary Material online), representing the largest sample size and population sampling of X-chromosomes in any nonmodel mammal. We compared intra- and interpopulation genetic diversity on the X-chromosome and autosomes. We estimated the ratios of X-to-autosome (X/A) nucleotide diversities (e.g., absolute and relative X/A diversity; Arbiza et al. 2014), and evaluated their patterns as a function of distance from genes. We also compared these patterns in domestic populations of different geographic origin and between *Ovis* species. Further, we investigated potential explanations for the observed patterns by examining the effects of selection, sex-biased migration, admixture, and genetic drift. Our goals here were 3-fold. First, we aimed to detect general patterns of X-chromosome versus autosomes diversity among species as well as among domestic populations during domestication, demographic expansions, and breeding processes. Second, we sought to disentangle these different forces in shaping the patterns of X-chromosome evolution. Third, we examined selective footprints on the X-chromosome and related these signals to early domestication as well as recent artificial selection in breeding practices and natural selection in different environments.

Materials and Methods

The Ovine SNP50K BeadChip and Whole-Genome Sequence Data

We used publically available Illumina Ovine 50 K (54,241 SNPs) BeadChip SNP and whole-genome sequence data of domestic sheep (*O. aries*) and their wild relatives (e.g., *O. orientalis*, *O. ammon*, *O. musimon*, *O. vignei*, *O. dalli*, and *O. canadensis*). The ovine SNP BeadChip data consisted of genotypes of 943 females representing 68 worldwide domestic populations/breeds in previous investigations (Kijas et al. 2012; Zhao et al. 2017), in addition to 6 males and 2 females of *O. orientalis* obtained from the NextGen Project (http://projects.ensembl.org/nextgen/) (fig. 1 and supplementary table S1, Supplementary Material online). Details on quality control for the Ovine SNP50K BeadChip genotypes were provided in Supplementary Material.

In addition, we obtained a total of 110 whole genome sequences of domestic and wild sheep with a sequencing depth of 5–14× (fig. 1 and supplementary table S2, Supplementary Material online). Among the genomes, sequences of 63 domestic and 26 wild sheep (17 *O. orientalis*, 4 *O. vignei*, 2 *O. dalli*, and 3 *O. canadensis*) were from the NextGen Project, and 19 additional domestic and wild (1 *O. ammon* and 1 *O. musimon*) sheep sequences were from our previous work (Yang et al. 2016) (fig. 1 and supplementary table S2, Supplementary Material online). Quality control, filtering, and alignment for sequencing reads were detailed in Supplementary Material.

SNP Calling and Validation

We conducted population-scale SNP calling using a Bayesian approach as implemented in the SAMtools "mpileup" command (Li et al. 2009) with the parameters "*-pup -q 1 -C 50 -S -D -m 2 -F 0.002*" "*-popf –Q 20 –d 1 –D 10000000*." Only the bases with quality scores >20 ($Q \geq 20$) were considered, and then we kept the positions in which genotypes are reported from at least one individual for the nucleotide diversity calculation. After SNP calling, we filtered raw variants at the individual level using the following criteria, and SNPs meeting either of the criteria were removed: 1) the SNPs are not in biallelic sites; 2) low-quality SNPs with the coverage depth < 5 or > 100 and RMS (root mean square) mapping quality < 20. To minimize the number of false positive calls, we also applied more strict filtering criteria in domestic and wild sheep, separately: 1) MAF < 0.05; 2) the proportion of missing genotypes > 50%; and 3) *P* value of the chi-squared test for Hardy–Weinberg equilibrium (HWE) test > 0.00001 (*P* value > 0.00001). SNPs meet any of the above criteria were filtered. These high-quality SNPs were referred as whole-genome sequence (WGS) SNPs.

Further, we examine the accuracy of our high-quality SNPs using the method as described in Yang et al. (2016). In summary, we compared the SNPs with those from Build 143 of the sheep dbSNP database in the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/SNP). We annotated all the high-quality SNPs using the ANNOVAR software (Wang et al. 2010) and categorized them into three groups according to their genomic locations such as intronic, exonic, and intergenic.

Estimates of Nucleotide Diversity and Its Pattern along X Chromosome

We calculated within-population genomic diversity metrics and estimated the linkage disequilibrium (LD) pattern between pairwise SNPs (Supplementary Material online). In addition, genetic differentiation and population structure were examined based on the Ovine BeadChip and WGS data sets (Supplementary Material online).

The nucleotide diversity (π) on per-site basis was calculated using the formula by Nei and Li (1979) in VCFtools (Danecek et al. 2011). For SNPs in intronic, exonic, and intergenic locations, we calculated π separately. To correct for the effect of potential difference in mutation rates between X-chromosome and autosomes, we calculated mean interspecies divergence (*D*) (Jukes and Cantor 1969) for X-chromosome and autosomes between sheep and cattle (*Bos taurus*) using the software MEGAv.7 (Kumar et al. 2016) under the Jukes–Cantor model (Jukes and Cantor 1969). We then normalized the nucleotide diversity (π) using the *D* estimates following the method in Gottipati et al. (2011). In summary, we aligned sequences of domestic sheep and cattle from UCSC (http://hgdownload.soe.ucsc.edu/goldenPath/oviAri1/vsBosTau4/), and computed the mean sheep–cattle divergence as the fraction of fixed differences using a Jukes–Cantor correction for recurrent mutations (Jukes and Cantor 1969).

In order to examine the impact of selection on linked sites via affecting estimates of nucleotide diversity and absolute X/A diversity (X-to-autosome nucleotide diversity ratio), we evaluated the above two diversity indices as a function of physical distances from the nearest genes in 1 Mb windows, and sorted these windows into 20 bins at an increasing distance. Physical distance from each SNP site to the nearest gene was estimated following the method of Nam et al. (2017).The 95% confidence intervals for the estimates of π and absolute X/A diversity were obtained using a nonparametric bootstrapping procedure of resampling 1000 random data set from each bin.

Further, we estimated relative X/A diversity, the ratio of the estimates of absolute X/A diversity (Gottipati et al. 2011), to unravel differential forces of demographic history and selection among populations from different continental origins (e.g., Middle East, Central and East Asia, South Asia, Africa, America, and Europe). Different from the absolute X/A diversity, relative X/A diversity overcomes the effects stemming from differences in mutation rates between X-chromosome

and the autosomes and other potential issues (e.g., difference in SNP ascertainment biases between X-chromosome and autosomes) (Arbiza et al. 2014). We assumed the initial sheep domestication in the Middle East (Zeder 2008; Lv et al. 2015), and calculated the relative X/A diversity between non-Middle Eastern populations to the Middle Eastern populations following the method of Gottipati et al. (2011). We also examine the relationship between relative X/A diversity and physical distances from the closest gene in the comparisons as detailed above.

Analysis of Sex-Biased Migration

To explore the potential effect of sex-biased demographic events on the observed pattern (Keinan and Reich 2010), we calculated Q statistics, a test for sex-biased migration (Goldberg et al. 2017). Under a simple demographic model, that is, two isolated populations originated from one common ancestor, and have been panmictic and constant in population size since their split, autosome-to-X genetic drift ratio (Q) is equal to 0.75 (Keinan et al. 2009). We measured sex-biased genetic differentiation on autosomes and X chromosome, that is, autosomal (F_{ST}^{A}) and X-chromosomal (F_{ST}^{X}) F_{ST} values, and then estimated the Q statistic in the formula $Q = \ln(1 - 2F_{ST}^{A})/\ln(1 - 2F_{ST}^{X})$, following the methods of Keinan and Reich (2010) and Waldman et al. (2016). Based on the assumption of initial sheep domestication in the Middle East (Zeder 2008; Lv et al. 2015), we calculated the Q statistic between different continental groupings using the X-linked and autosomal SNPs. From the theoretical prediction, the Q statistic is equal to 0.75, while deviation from the expected value suggests possible occurrence of sex-biased demographic events (Ramachandran et al. 2008; Emery et al. 2010).

Simulations of Genetic Drift

We simulate the genetic drift on X-chromosome and autosomes in the WGS data using an updated version of the POWSIM software (Ryman and Palm 2006) as detailed in Lamichhaney et al. (2012) and Lamichhaney et al. (2017). Simulations of the expected distribution of F_{ST} among domestic populations were conducted under a selective neutrality model. We then compare the genetic drift on X and autosomes using the distributions of simulated and observed F_{ST} values. To obtain the selectively neutral SNP loci, the following criteria were applied, and SNPs met any of the criteria were removed: 1) SNPs are within genic regions; 2) SNPs had less than five reads from each sample; 3) SNPs had >20% missing rates in a population; 4) SNPs had negative observed F_{ST} values. After the filtering, we obtained 27,167 autosomal and 1,795 X-linked neutral SNPs. For X-linked SNPs, we simulated an infinitely large base population with a major frequency of 0.863, and the base population was split into six subpopulations. The simulation was run for 1,795 repetitions with the effective size of $N_e = 8{,}350$, generation time of $t = 840$ (3 years/generation), and expected $F_{ST} = 0.049$. For autosomal SNPs, simulation with 27,167 repetitions was conducted with a major frequency of 0.899 for the base population and six subpopulations. The simulation was run with the effective sizes of $N_e = 10{,}000$, generation time of $t = 840$ (3 years/generation), and expected $F_{ST} = 0.0412$.

Tests for Selective Sweeps on the X-Chromosome and Autosomes

To characterize the selective forces shaping patterns of genetic variability on the X chromosome and autosomes, we applied three statistical tests to detect signatures of selective sweeps. We calculated F_{ST} values (Weir and Cockerham 1984), integrated haplotype scores (*iHS*) (Voight et al. 2006), and cross-population extended haplotype homozygosity scores (*XP-EHH*) (Sabeti et al. 2007) between domestic sheep and their putative wild ancestors Asian mouflon as well as between different continental groups of sheep populations.

In the F_{ST}-based test, we calculated the average F_{ST} value for 1 Mb nonoverlapping windows based on the Beadchip using GENEPOP v4.3 (Rousset 2008), and for 200-kb nonoverlapping windows based on the WGS SNPs using VCFtools (Danecek et al. 2011). In the *iHS* and *XP-EHH* analyses, we first inferred ancestral alleles by comparing the BeadChip and the WGS SNPs with sheep.oarv3_1.ancestral_alleles (BeadChip SNPs: http://www.sheephapmap.org/download/results/AncestralAlleles; WGS SNPs: ftp://ftp.ebi.ac.uk/pub/databases/nextgen/ovis/variants/supplemtary_info/; Jiang et al. 2014). Only SNPs which possess ancestral alleles (BeadChip: 1,132 X-chromosomal and 44,031 autosomal SNPs; WGS: 52,325 X-chromosomal and 6,073,533 autosomal SNPs) were kept for the following analyses. We then inferred haplotypes for each sample using the SHAPEIT v2 program (Delaneau et al. 2013). Finally, we computed the fraction of extreme *iHS* ($|iHS| > 2$) or *XP-EHH* ($XP\text{-}EHH > 2$) and its empirical P values for each pair of comparison using 1 Mb and 200-kb nonoverlapping windows for the BeadChip and the WGS SNPs, respectively. Both the analyses were implemented in Selscan (Szpiech and Hernandez 2014) assuming 1 Mb = 1 cM across the sheep genome. In addition, we quantified and compared the selective intensities between X-chromosome and autosomes by estimating the ratio of selective SNPs (i.e., $|iHS| > 2$ or $XP\text{-}EHH > 2$) to all the X-chromosomal or autosomal SNPs.

Further, we conducted functional annotation for genes in selective regions using Oar_v4.0 (https://www.ncbi.nlm.nih.gov/genome/?term=sheep). For those candidate genes, we performed the phenotype enrichment analysis in the MGI database (Mouse Genome Informatics; http://www.informatics.jax.org/) using the Enrichr program (Chen et al. 2013; Kuleshov et al. 2016). In addition, we implemented Gene Ontology (GO) enrichment and pathway analyses using the

DAVID Bioinformatics Resources 6.8 (available from http://david.abcc.ncifcrf.gov) (Huang et al. 2009a, 2009b). We set all known sheep genes as the background gene set and assessed the statistical significance (*P* value) by the Binomial distribution test.

Results and Discussion

Genome Assembly and SNPs

We analyzed a total of 3,760 Gb raw pair-ended sequence reads from 110 whole genomes (82 domestic and 28 wild individuals) from seven *Ovis* species at an average coverage of$\sim$ 8.4$\times$ for both X-chromosome and autosomes (supplementary table S3, Supplementary Material online). After strict filtering, we obtained a total of 11,743,108 (autosome: 11,420,115; X: 322,993) high-quality SNPs in sheep and 28,926,374 (autosome: 28,358,469; X: 567,905) high-quality SNPs in their wild relatives (supplementary table S4, Supplementary Material online), with 7,605,294 SNPs shared between domestic and wild species. As expected, most of the high-quality SNPs were located in intergenic regions, followed by intronic and exonic regions (supplementary table S5, Supplementary Material online). Overall, 94.43% of the SNPs identified in sheep and 79.76% in wild ovines were validated using the sheep dbSNP database, indicating a high reliability of the called SNP variants (SNP validation rate in sheep: 89.46–96.21%; Yang et al. 2016; compared with, for examples, cross-species SNP validation for domestic horse *Equus caballus* in the Przewalski's horse *Equus ferus przewalskii*: 50.8%; McCue et al. 2012).

For the WGS data, 280,304 and 442,000 X-linked SNPs in domestic and wild sheep were analyzed, excluding those in the pseudoautosomal region (PAR: 42,689 and 125,905 SNPs in domestic and wild species, respectively; see the Results below). After the quality control and removing SNPs in the PAR (7 SNPs; see below), the BeadChip data set consisted of 45,822 autosomal and 1,187 X-linked SNPs, from a wide collection of native sheep breeds from around the world with varied phenotypic traits such as the types of main product (e.g., wool, meat, and milk), growth rate, body size, tail type, coat color, and disease resistance.

X-Chromosome Pseudoautosomal Region

To locate the PAR, the observed heterozygosity (H_O) for each SNP in the BeadChip was calculated within females and an additional set of 924 males from the world's sheep populations, separately. In the BeadChip, the first 7 SNPs, located between the positions 1.12–5.55 Mb (SNPs: rs400492871–rs401737908), showed an average $H_O = 0.25$ in males (supplementary fig. S1, Supplementary Material online), whereas almost all the remainder 7.14–134.93 Mb showed estimates of $Ho = 0$ (supplementary fig. S1*B*, Supplementary Material online). In contrast, an average estimate of $H_O = 0.31$ was observed for those in females (supplementary fig. S1*A*, Supplementary Material online). Therefore, these observation suggested that the PAR is likely located between 0 and 7.10 Mb on the X-chromosome in sheep. This observation is roughly congruent with the recombining region (i.e., PAR) in the sheep reference genome assembly (Jiang et al. 2014), in which the genomic region 0–7.05 Mb was suggested for the PAR.

In the WGS data, we observed significantly lower (Wilcoxon $P < 0.01$) mean π estimates in both the autosomes and X-linked region in sheep ($\pi_A = 0.0031$, $\pi_X = 0.0016$) than counterpart values in Asian mouflon ($\pi_A = 0.0038$, $\pi_X = 0.0019$) (fig. 2 and supplementary table S6, Supplementary Material online), which is consistent with a loss of genetic diversity due to a population bottleneck during domestication (VanBuren et al. 2016). In both sheep and Asian mouflon, we also observed a pronounced shift in π on both sides of the PAR boundary: average estimate of π in the PAR was $\sim$3.6-fold higher than in X-linked regions (fig. 2*A* and *B* and supplementary table S6, Supplementary Material online). This observation could be explained by a higher recombination rate in the PAR than the X-specific portion during male meiosis (Galtier 2004; White et al. 2012). Similar patterns of variation in π were also observed in all the other wild sheep (supplementary table S6, Supplementary Material online). Taken together previous inferences about the location of the PAR in sheep and other ruminants (Van Laere et al. 2008; Das et al. 2009; Jiang et al. 2014;), patterns of both genetic diversity indices (i.e., H_O and π) supported for the first time a clear and consistent PAR boundary ($\sim$7.10 Mb) across the genus *Ovis*.

We observed significantly lower (Wilcoxon $P < 0.01$) genetic divergence (i.e., mean F_{ST} values across loci) in the PAR region than the X-linked regions between sheep populations as well as between sheep and Asian mouflon (fig. 2*C* and *D*). A similar pattern of lower mean F_{ST} in the PAR than the X-linked regions has also been observed among worldwide pig populations and between domestic pigs and wild boar (Burgos-Paz et al. 2013). Similar patterns of genetic differentiation for PAR versus X-linked between domestic animals and their wild species could be partially attributed to some early genetic introgression (e.g., via interbreeding) from the wild relatives into domestic species following domestication, which is known to have occurred commonly (Larson and Burger 2013; Lv et al. 2015; Barbato et al. 2017; Zhao et al. 2017). However, an elevated divergence in PAR relative the X-linked region has been found between human and other primate species such as chimpanzee and macaque, potentially owing to complex speciation events (Patterson et al. 2006; Cotter et al. 2016).

Within-Population Genetic Diversity and between-Population Genetic Differentiation for X-Chromosome versus Autosome

In general, within-population genetic diversity indices (H_O, H_E, P_n, and N_e) estimated from X-linked SNPs were lower than

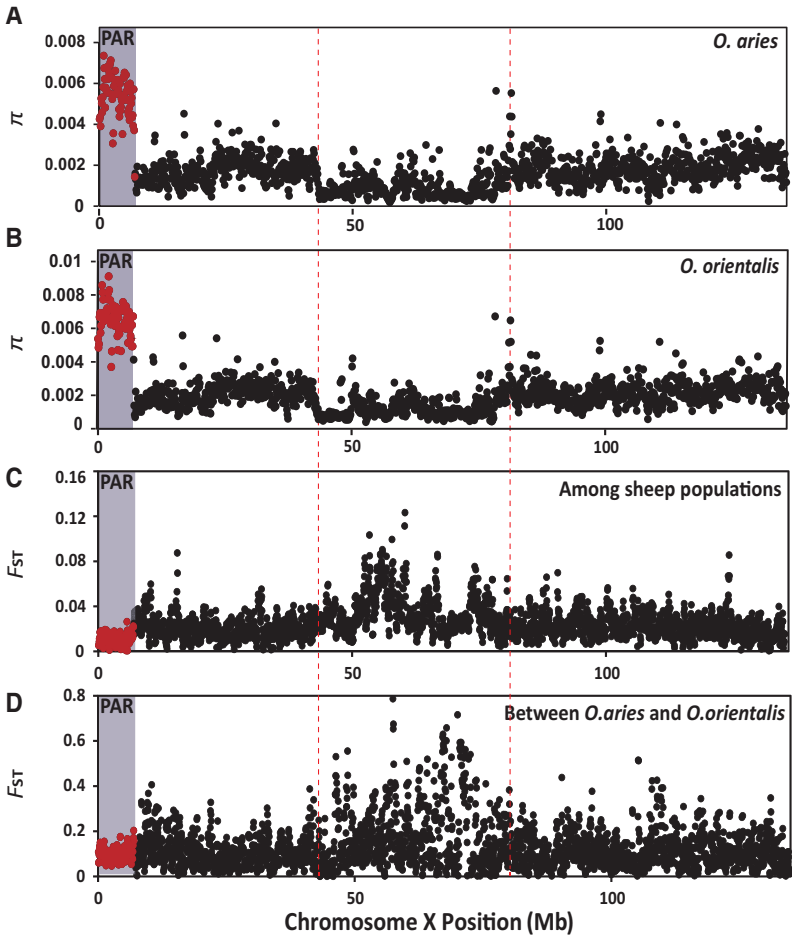

Fig. 2.—Nucleotide diversity and divergence along the X chromosome. (*A*) Nucleotide diversity (π) values of domestic sheep (*Ovis aries*) are showed in 50 kb windows and 10 kb steps along chromosome X. (*B*) Nucleotide diversity (π) values of Asian mouflon (*O. orientalis*) are showed in 50 kb windows and 10 kb steps along chromosome X. (*C*) F_{ST} dynamic variation among domestic sheep (*O. aries*) populations are plotted in 100 kb windows and 50 kb steps along chromosome X. (*D*) F_{ST} dynamic variation between sheep (*O. aries*) and Asian mouflon (*O. orientalis*) are plotted in 100 kb windows and 50 kb steps along the X chromosome. The pseudoautosomal region (PAR) is represented in red dots. A sequential and conserved domain with lower nucleotide diversity and higher divergence is marked by the red dotted lines.

those from autosomal SNPs in the BeadChip data (supplementary table S7, Supplementary Material online). Meanwhile, we observed higher level of LD in the X-linked region than chromosome 6, similar in the size and the number of SNPs to those in the X-linked region, within populations (supplementary fig. S2, Supplementary Material online). Additionally, the X-linked region showed more ROHs, a larger average size of ROHs and a higher coverage of ROHs than autosomes in sheep and the wild species (supplementary figs. S3 and S4, Supplementary Material online). The BeadChip data showed a reduction in genetic diversity for the X-linked versus autosomal SNPs at a population and continental level, but this was much smaller than the observed reduction in π of the X-linked region relative to autosomes in the WGS data (supplementary tables S6 and S7, Supplementary Material online) and the 1/4 reduction under assumption of equilibrium (Charlesworth et al. 1987; Ellegren 2009). Similar differences have been observed in X chromosome analyses of worldwide pig populations using SNP and sequence data (Amaral et al. 2011; Esteve-Codina et al. 2011; Burgos-Paz et al. 2013). The most likely explanation for our results is SNP ascertainment bias in the BeadChip as indicated by the longer branch lengths in the phylogenetic trees built based on the BeadChip data than those based on the WGS data (Supplementary Material and supplementary fig. S5, Supplementary Material online), in which SNPs with the highest minor allele frequencies (MAFs) were selected for both the autosomes and the X-chromosome.

In the analyses of population genetic structure using BeadChip SNPs, both the autosomal and X-linked SNPs revealed apparent geographic differentiation (i.e., Middle East, Central and East Asia, South Asia, Africa, America, and Europe) in the STRUCTURE (with the optimal $K = 6$; see Supplementary Material and supplementary figs. S6 and S7,

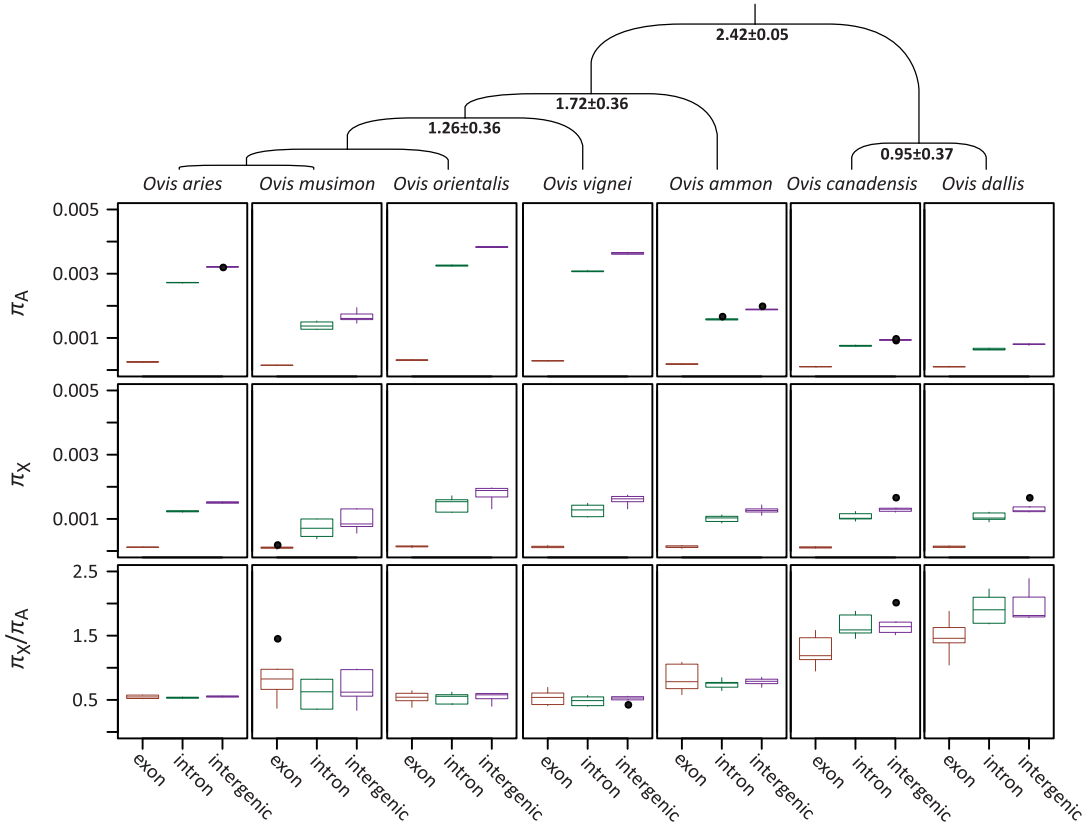

Fig. 3.—Nucleotide diversity (π) and the X-linked region/autosome diversity ratio (X/A) in exons, introns, and intergenic regions in the seven *Ovis* species. The phylogenetic relationships among the seven *Ovis* species investigated and the divergence times (in Ma) at the top were adapted from figure 1 in Rezaei et al. (2010) The top, middle (within the box), and bottom boundary lines of the boxes represent 25%, 50% (median value), and 75% of π.

Supplementary Material online) and AMOVA analyses (supplementary table S8, Supplementary Material online). Nevertheless, the X-linked SNPs showed significantly (Wilcoxon $P < 0.01$) higher percentage of molecular variance than autosomal SNPs between continental groups (X: 8.42%; autosome: 4.42%) as well as between populations within continental groups (X: 9.99%; autosome: 8.35%; supplementary table S8, Supplementary Material online). Similarly, higher levels of genetic divergence for X-linked SNPs relative to autosomal loci were observed in the WGS data (supplementary table S6, Supplementary Material online), while the SNPs in the PAR failed to separate the sheep population into continental groups in the PCA analyses (supplementary fig. S8, Supplementary Material online).

Patterns of Nucleotide Diversity along the X-Chromosome

In *Ovis* species, we observed the lowest mean π in exons followed by intronic and intergenic locations in the X-linked regions and autosomes (fig. 3). We found one conserved domain (43.05–79.25 Mb) featuring lower π in the X-linked region across all species (fig. 4). Nam et al. (2015) proposed three potential forces for the extremely low X-chromosomal nucleotide diversity in great apes, such as close inbreeding, selective sweeps, and low mutation rates in large genomic regions.

Here, inbreeding is unlikely to explain the pattern because no conserved region of low diversity was observed on the autosomes (supplementary fig. S9, Supplementary Material online). In order to clarify the selective effect in shaping these genomic variations, we identified selected genomic regions using the criteria of |*iHS*| (integrated haplotype score; Voight et al. 2006) > 2 and XP-EHH (cross-population extended haplotype homozogysity; Sabeti et al. 2007) > 2 in tests between domestic sheep and Asian mouflon, and further compared selection intensities between X-linked regions and autosomes. We found significantly ($P < 0.01$) larger numbers of selective SNPs per Mb in autosomes (|*iHS*|: 106.99 SNPs/Mb; *XP-EHH*: 87.72 SNPs/Mb) than in the X-linked region (|*iHS*|: 16.74 SNPs/Mb; *XP-EHH*: 12.65 SNPs/Mb) based on the WGS data (table 1). In addition, similar selective sweeps analyses in the BeadChip data showed lower intensities in the X-linked region than in autosomes (table 1). Different from previous evidence for stronger selective sweeps on X-chromosome than autosomes in the primates (Hammer et al. 2010; Nam et al. 2015), less intensive selection on the X-linked region observed here

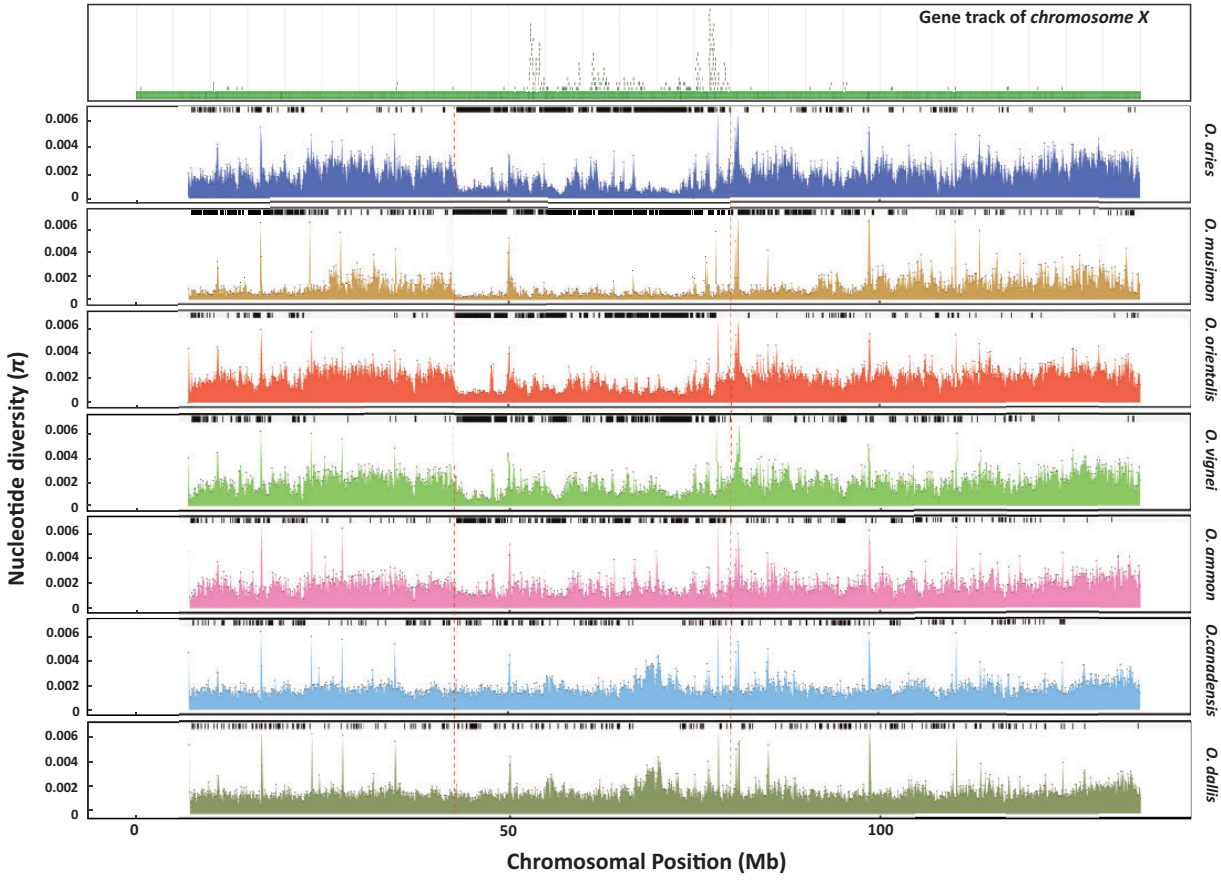

Fig. 4.—Nucleotide diversity along the X-linked regions for the seven species of *Ovis*. Black bars at the top of each panel indicate the regions where the nucleotide diversity is 20% less than the mean diversity across chromosome X in the species. A sequential and conserved domain with low nucleotide diversity level is marked by the red dotted lines. Distribution of protein coding genes on chromosome X was plotted at the top.

Table 1

Comparison of Selection Intensity (number of SNPs/Mb) between the X-Linked Region and Autosomes Based on the Selection Tests between Domestic Sheep and Asian Mouflon using the *iHS* (Voight et al. 2006) and *XP-EHH* (Sabeti et al. 2007) Approaches

	Illumina Ovine 50K BeadChip		Whole-Genome Sequences	
	iHS[a]	*XP-EHH*[b]	*iHS*[a]	*XP-EHH*[b]
X-linked region	0.34	0.22	16.74	12.65
Autosome	0.83	0.38	106.99	87.72

[a]Number of SNPs selected based on the threshold of $|iHS| > 2$.
[b]Number of SNPs selected based on the threshold of $XP\text{-}EHH > 2$.

could be due to the fact that the majority of QTLs (https://www.animalgenome.org/cgi-bin/QTLdb/OA/summary?summ=chro&qtl=1, 658&pub=132&trait=225) (supplementary fig. S10*A*, Supplementary Material online) and genes (supplementary fig. S10*B*, Supplementary Material online) (Xu and Li 2017) for production and functional traits under strong and long-term directional selection are located on the autosomes instead of the X-chromosome in domestic animals (supplementary fig. S10, Supplementary Material online). However, we did not detect a large conserved region of highly reduced nucleotide diversity on the autosomes. Thus, these results suggested that selective sweeps, at least alone, cannot explain the observation. In addition, this domain showed higher levels of genetic divergence between sheep and Asian mouflon (fig. 2*D*). Together with the evidence of low π values in the conserved domain in the ancestral species of sheep, our evidence implied that this pattern could not be ascribed to reduced recombination rates (Nam et al. 2015). Interestingly, we observed an extremely high density of coding genes in this

domain (fig. 4), and, thus, a large number of conserved coding regions in these genes could partly account for the large diversity reduction observed in the X-linked region.

In five of the species (sheep, Asian mouflon, European mouflon, argali, and urial), mean estimates of π in the X-linked regions was ~60% of those in autosomes for the exonic, intronic, and intergenic locations (fig. 3), which is substantially lower than the expected X/A ratio of 3/4 under neutral evolutionary models (Charlesworth et al. 1987). Nevertheless, the ratios were in the range (X/A ratio = 0.32–1.47) observed in previous studies in primates and other model species (Ellegren 2009; Hammer et al. 2010; Arbiza et al. 2014; Nam et al. 2015). Higher X/A diversity ratios (> 1) were found in *O. canadensis* and *O. dallis*, which are more distantly related to domestic sheep (fig. 3). Similar cases have rarely been observed in other species or populations except for African populations of *D. melanogaster* reported by Andolfatto (2001). Andolfatto's argument is that autosomal inversions work as recombination suppressors and will therefore lead to reduced autosomal variability when backgrounds selection or the frequency of selective sweeps is large (Andolfatto 2001). Mammals usually do not have the genomes littered with inversion like *Drosophila*, so it is an unlikely explanation. Here, the most likely explanation for the observation could be that the species are more divergent from domestic sheep and, thus, we can only map reliably in the least variable regions.

Patterns of Diversity Related to Distances from Genic Regions

In all the six continental groups of populations, the SNP estimates of π increased with physical distance from genes (in 1 Mb windows and the windows were sorted into 20 bins at an increasing distances) in autosomes and X-linked regions (fig. 5*A*), consistent with increased levels of selection on genes and linked regions (McVicker et al. 2009; Arbiza et al. 2014). However, a much lower and flatter trend in π values was observed at SNPs within 250 kb from genes in X-linked regions ($\pi \leq 0.002$) compared with autosomes ($\pi \leq 0.004$) (fig. 5*A*). In addition, we observed a mean normalized X/A ratio (0.58) that was significantly lower than the expected values of 0.75 (3/4). The X/A ratio showed a slight decrease with an increase in distance in all the geographic groupings (fig. 5*A*). This differs from a positive correlation between X/A diversity and distance to genes in human (McVicker et al. 2009; Gottipati et al. 2011; Arbiza et al. 2014) and the great apes (Nam et al. 2015; Narang and Wilson Sayres 2016), which has been mainly explained by stronger directional selection at genes and their linked sites on the X chromosome (Arbiza et al. 2014; Nam et al. 2015). Therefore, the lower impact on diversity-reduction in the X-linked region than the autosomes could be due to the more intensive recent directional selection on autosomes than in the X-linked region inferred above. An alternative explanation is the evolutionary inertness of the X-chromosome, which by the hemizygosity in male and the X-inactivation in females are subject to an intensive purifying selection and thus do not follow the normal evolutionary changes as observed for autosomes.

As in domestic sheep, we also observed similar overall trends of π and X/A diversity with distance from genes in wild sheep (fig. 5*B*). In particular, we found that π was nearly constant with the distance from genes on the autosomes (*O. canadensis*: $\pi = 0.00127$–0.00145; *O. dallis*: $\pi = 0.00115$–0.00123), while π decreased with distance from genes in the X-linked regions (fig. 5*B*) in *O. canadensis* and *O. dallis*. In addition, the X/A diversity ratio (*O. canadensis*: 1.29–1.75; *O. dallis*: 1.47–1.99) were >1 and showed a decrease with distances from genes in *O. canadensis* and *O. dallis* (fig. 5*B*). Similar patterns have rarely been revealed in primates and other model species (but see Evans et al. 2014). Forces including sex differences in reproductive success (Charlesworth 2012), varying population size and natural selection on GC content (Evans et al. 2014) have been invoked to explain such observations in different scenarios. However, we cannot differentiate the relative contribution of difference forces to the pattern observed here due to the limitation of data (e.g., the field data and de novo genomic data) in *O. canadensis* and *O. dallis*.

Relative X/A Diversity in Continental Population Groupings

In all five pairwise comparisons based on the WGS data, relative autosomal diversity was nearly constant within 250 kb distance from genes (fig. 6*A*). Contrasting with the general assumption of higher diversity at the domestication center, we found significantly higher (~3%) autosomal diversity in African, American, and European than in the Middle Eastern populations (fig. 6*A*). A possible explanation could be multiple introgression from the Middle Eastern domestication center and admixture in these regions (Chessa et al. 2009). In addition, we noted that the Middle Eastern populations are fat-tailed and probably do not represent wholly the original thin-tailed domesticates that spread to Europe, Africa, and America. For the five pairwise groups, the relative X-linked diversity was approximately equal between the Middle Eastern and the Central and East Asian populations within 250 kb from the genes (fig. 6*B*), with no significant correlation with distance from genes (the Pearson's $r = 0.45$, one-tailed test $P = 0.132$). In contrast, we observed slight lower X-linked diversity in the Middle Eastern populations than populations of the other four continental groupings (fig. 6*B*), and the relative X-linked diversity showed a significant positive correlation (the Pearson's $r = 0.711$–0.925, one-tailed $P < 0.05$) with distance from genes for South Asian and African populations within 250 kb.

We also estimated relative X/A diversity among different continents groups. Four pairs of the relative X/A diversity (SA/

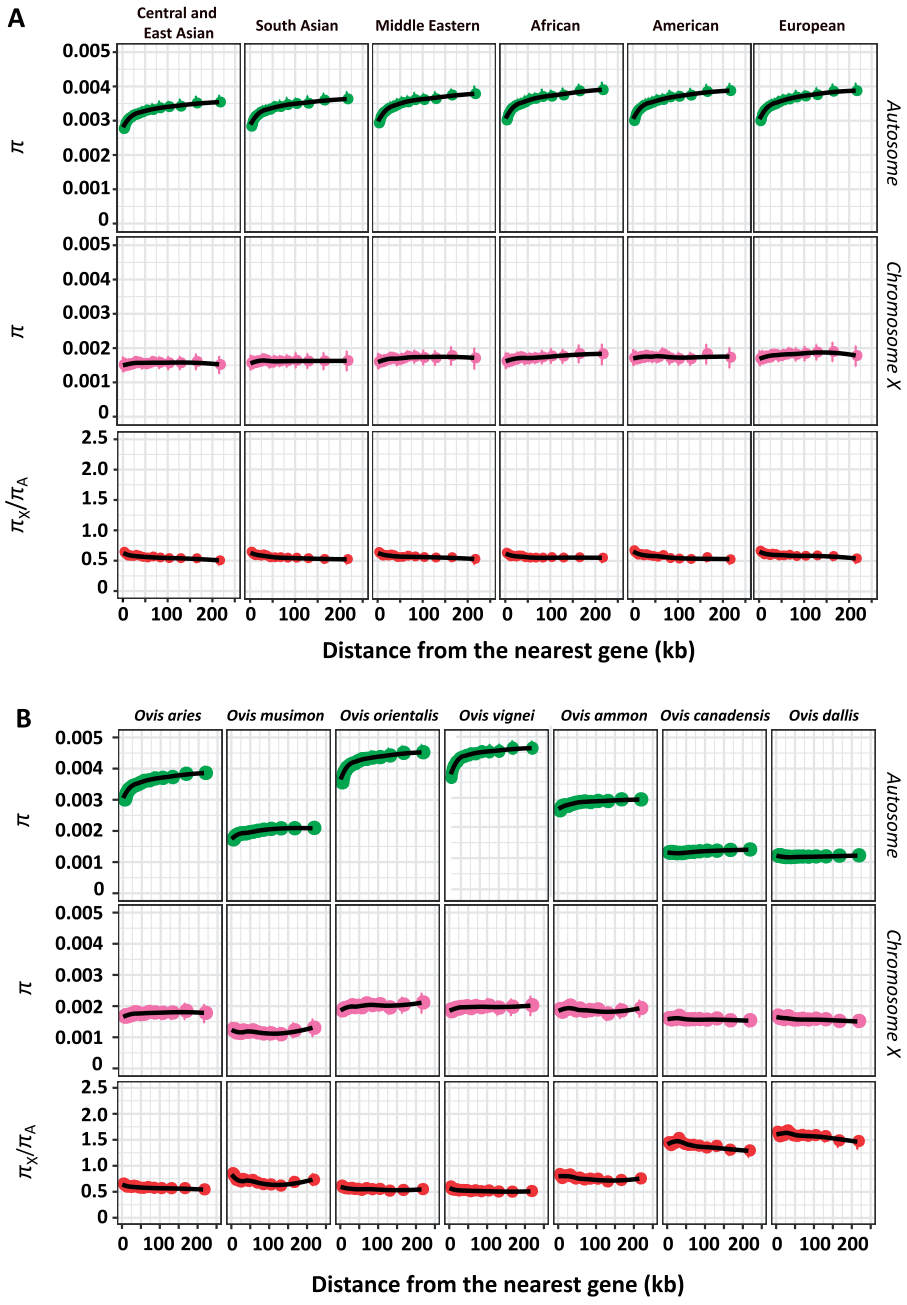

Fig. 5.—Nucleotide diversity (π) and the X-chromosome/autosome diversity ratio (π_X/π_A) with the physical distance to the nearest gene in the *Ovis* species. (*A*) π and π_X/π_A in the continental groupings of domestic populations. (*B*) π and π_X/π_A in the seven *Ovis* species. 95% confidence intervals from 1,000 bootstrapping iterations are showed in each bin.

ME, AF/ME, AM/ME, and EU/ME) showed significant positive (the Pearson's $r = 0.61$–0.77, one-tailed $P < 0.05$) correlation with distances from genes (fig. 6*C*), which could be due to the moderate differences in X-linked diversity among populations from different continents groups. Remarkably, within the domestic sheep the X-linked diversity is more variable across different continents than autosomal diversity (fig. 6*A* and *B*; supplementary tables S6, S8, and S9, Supplementary Material online). It may be noted that purifying selection of mtDNA diversity has been postulated to act on the long term (Ho et al. 2011), and, thus, similar long-term purifying selection might have acted on the X-linked region as well. The remaining X/A comparison (CEA/ME) showed a lack of correlation with distances from genes (the Pearson's $r = 0.206$, one-tailed $P = 0.312$). The relative X/A diversity between Central and East Asian and the Middle Eastern populations was nearly constant (X/A = 1.043) across all the scales of distances (≤ 250 kb) from genes (fig. 6*C*), which indicated genomic

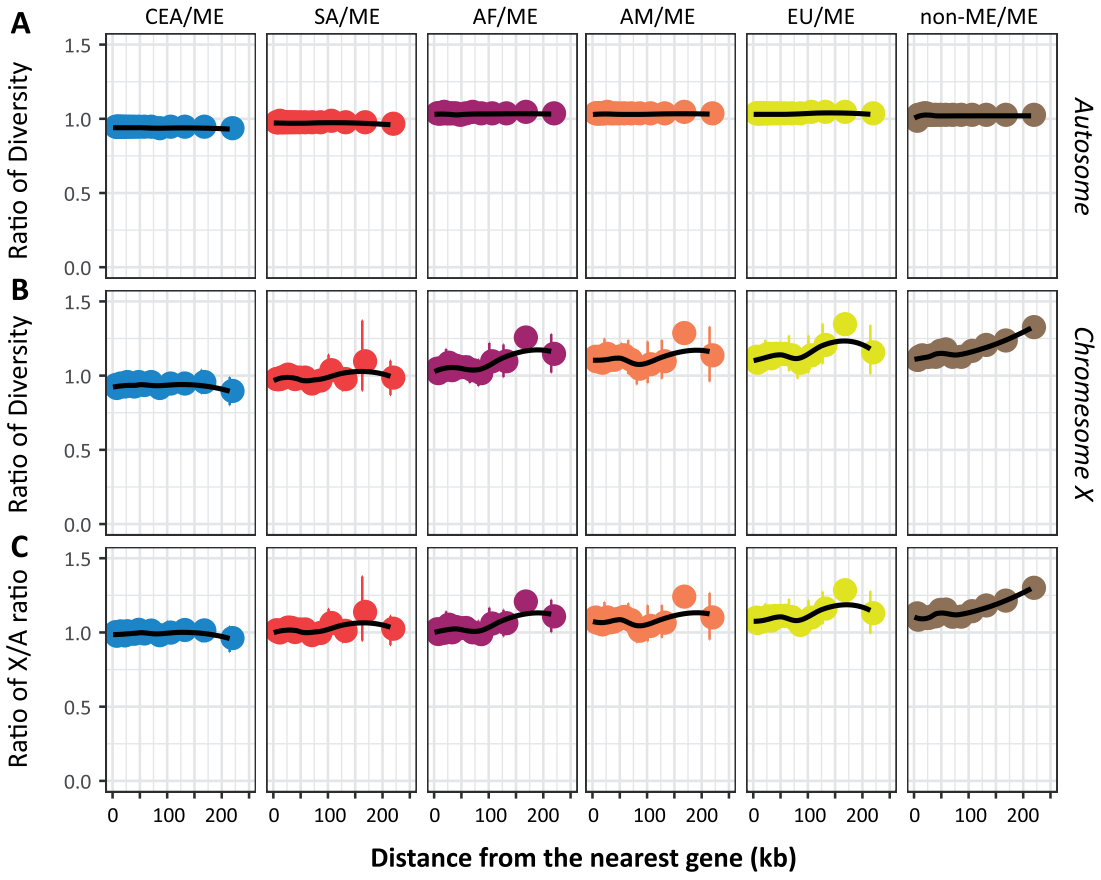

Fig. 6.—Relative nucleotide diversity (π) and their ratio (X/A) with the physical distance to the nearest gene for the six continental groupings of sheep populations. (*A*) The ratio of normalized diversity in the groupings of populations from the Middle East (ME) domestication center to those in the other five groupings of populations with the physical distance to the nearest gene for X-linked region. (*B*) The ratio of normalized diversity in the groupings of populations from the Middle East (ME) domestication to those in the other five groupings of populations with the physical distance to the nearest gene for autosome. (*C*) Relative X/A diversity in each pair groupings of populations with the physical distance to the nearest gene. 95% confidence intervals from 1,000 bootstrapping iterations are showed in each bin. CEA, Central, and East Asian populations; ME, Middle Eastern populations; SA, South Asian populations; AF, African populations; AM, American populations; EU, European populations. Non-ME, Central and East Asian/South Asian/African/American/European populations.

similarity and similar effect of gene proximity between populations from the two regions. This is most likely due to frequent genetic exchange between populations from the two geographic groups in the past (Zhao et al. 2017). In addition, we observed nearly constant pattern of autosomal diversity, but apparently increasing pattern of X-linked diversity and relative X/A diversity with increased genetic distances for the comparison between Middle Eastern and non-Middle Eastern populations (fig. 6*C*). In human, an increase in both autosomal and X-linked nucleotide diversity, but an invariant X/A diversity, with distance from genes was observed in the comparison between African and non-African populations (Arbiza et al. 2014). The difference in these patterns between sheep and human could be partly due to different demographic forces acted on them.

In addition, we noted that most of the genome sequences have the coverage of 10–20× and, consequently low or rare variants were missed. In addition, the number of false calls, despite the statistical treatments and quality control (QC) filtering, might be relatively higher. Thus, the allele callings might influence the final estimates. Nevertheless, we ran a site frequency spectrum (SFS) using the program ANGSD v.0.902 (Analysis of Next Generation Sequencing Data; Korneliussen et al. 2014) and did not find significant (Student's *t*-test $p = 0.503$) difference in allele frequencies between the ones observed and those expected under the mutation-drift equilibrium (supplementary fig. S11, Supplementary Material online). In this sense, the influence of missed low or rare variants and false calls could be very slight, if there were.

Genetic Drift on X and Autosomes

To further investigate whether the patterns of nucleotide diversity and X/A diversity related to distance to genes were

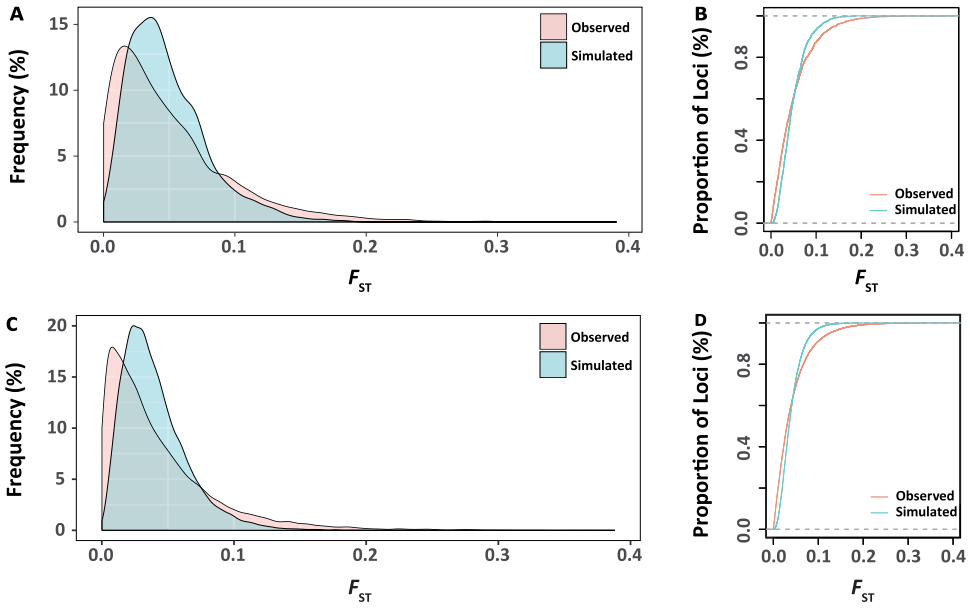

Fig. 7.—Distribution of observed and simulated F_{ST} values. (*A*) Density distribution of F_{ST} values in simulated and observed data sets on X chromosome. (*B*) Cumulative distribution of the observed and simulated F_{ST} values on X chromosome. (*C*) Density distribution of F_{ST} values in the simulated and observed data sets on autosome. (*D*) Cumulative distribution of observed and simulated F_{ST} values on autosome.

primarily driven by the genetic drift, we performed the simulations of genetic drift on X-chromosome and autosomes. For X chromosome, there was a good agreement between the observed and simulated F_{ST} data (fig. 7*A* and *B*). Around 60% of the loci, both in the simulated and observed data sets, showed low F_{ST} values in the range of 0–0.049 (fig. 7*B*). Under the selectively neutral model, average value of simulated F_{ST} (0.049) is roughly equal to the average value of observed F_{ST} (0.050). The slightly higher average value of observed F_{ST} was derived from a bit longer tail of F_{ST} values in the observed data set, which could be caused by selection (Lamichhaney et al. 2012). For autosomes, similar results were observed in the distributions of simulated and observed F_{ST} values (fig. 7*C* and *D*). As a result, comparison of the simulated distributions of F_{ST} values indicated an accelerated genetic drift on X chromosome (average $F_{ST} = 0.049$) than on autosome (average $F_{ST} = 0.040$). Thus, our results suggest that genetic drift is one of the primary forces shaping the differential genomic patterns on X-chromosome and autosomes. Nevertheless, only the effect of genetic drift, but not other complex demographic scenarios (e.g., domestication bottlenecks), has been considered in the coalescent simulation, Thus, caution is needed when interpreting the results, because it still cannot fully disentangle the effects of demography and reduced or weak selection.

Scan for Selective Signatures

In the selection tests between sheep and Asian mouflon using the *iHS* (integrated haplotype score), F_{ST}, and *XP-EHH* (cross-population extended haplotype homozogysity) statistics (Weir and Cockerham 1984; Voight et al. 2006; Sabeti et al. 2007) we identified a total of 54 and 45 extreme windows in the X-linked region and autosomes in the WGS data, respectively (supplementary fig. S12, Supplementary Material online). In these selective windows, we annotated 60 genes on the X-chromosome and 93 genes in the autosomes (supplementary table S10, Supplementary Material online). In the phenotype enrichment analysis using the MGI database, we found a set of functional genes associated with behavioral response to stress, nervous system development, lethality, pigmentation, and infertility (fig. 8). In particular, a number of selective genes were detected in the X-linked region related to the abnormal behavioral phenotypes, such as abnormal response to novelty, hypoactivity, reduced long-term potentiation, abnormal hair growth, and abnormal heart and ventricular development (fig. 8*A* and supplementary tables S11 and S12, Supplementary Material online). These phenotypes are the most likely traits that have been under strong selection during domestication (Kerje et al. 2004; Carneiro et al. 2014; Jensen 2014; Medugorac et al. 2017).

Overall, we detected four common regions across the three tests, and seven genes (*TRIM24*, *WDR36*, *CAMK4*, *STARD4*, *HAO1*, *ADRA1D*, and *SMOX*) were annotated (supplementary fig. S13, Supplementary Material online). Of these, three (*WDR36*, *CAMK4*, and *ADRA1D*) function in the behavior and nervous system. In particular, *CAMK4*, found in the neuropathic pain signaling pathway, affects nervous system development and function in yak and relates to reduced fear or increased stress activity (Medugorac et al. 2017). In addition, *ADRA1D* is enriched in the neuroactive

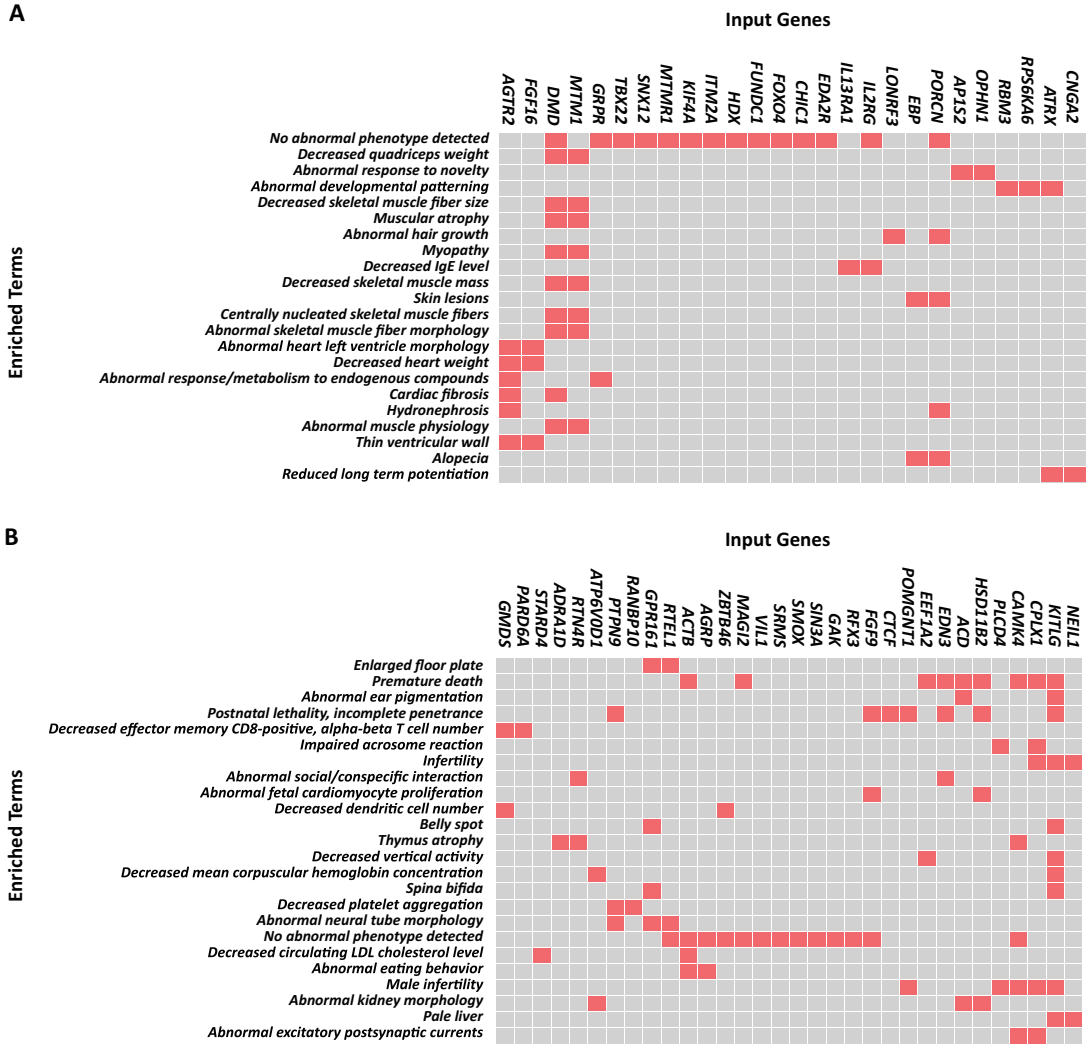

Fig. 8.—Significantly associated phenotypes for the selective genes in the phenotype enrichment analysis through the MGI database (Mouse Genome Informatics; http://www.informatics.jax.org/) using the Enrichr program. (*A*) The X-linked candidate genes. (*B*) The autosomal candidate genes.

ligand-receptor interaction pathway and potentially regulates abnormal behavior in the horse to adapt to new environments (Librado et al. 2015).

In order to detect genomic regions under recent selection, similar tests were also implemented between continental population groupings using the Beadchip data, which consisted of a worldwide collection of native sheep (supplementary table S1, Supplementary Material online). We detected a total of 72 selective windows in the comparisons between the five continental groupings and the Middle East populations (supplementary fig. S14, Supplementary Material online). Of these selective windows, 19 were overlapping in two of the three different tests: only two windows (73–74 Mb and 74–75 Mb) were detected in at least two different continental regions, while the rest 17 were in specific geographic regions (supplementary table S13, Supplementary Material online). In the two windows, we annotated four genes, and two of them (*DACH2* and *POF1B*) are related to the ovarian development, and could be under similar recent artificial selection for high fertility. We also annotated a total of 53 genes (supplementary table S13, Supplementary Material online) in the other seventeen windows identified in specific geographic regions. We found that a majority (64%, 34/53; e.g., *C1GALT1C1*, *IL1RAPL2*, and *LAMP2*) of genes are associated with metabolism, immunity, and diseases. Nevertheless, we did not found any enriched GO terms and pathways for the candidate genes, which could result from natural selection for local adaptation and resistance to various pathogens in different environments (Lv et al. 2014; Yang et al. 2016).

Conclusions

In conclusion, we present a comprehensive analysis of X-chromosome and autosomal diversity across the *Ovis*

genus. We extend previous efforts in quantifying the patterns of X-chromosome diversity and X/A ratio to domestic animals and their wild relatives. We observed greater reduction in X-linked diversity relative to autosomal diversity in these species. In contrast to the typically positive correlation observed in the primates and other model animals, we detected decreased X/A diversity as a function of distances from genes. Our evidence suggested that reduced directional selection but accelerated genetic drift in the X-linked region than autosomes as well as sex-biased migration could have led to the X/A diversity pattern. In addition, our results provide evidence that demographic events including multiple migrations and admixture during the expansion process are likely to have played a role in shaping patterns of the relative X/A diversity between the Middle Eastern and non- Middle Eastern populations. Finally, we identified a set of novel candidate targets of artificial and natural selection on the sheep X-chromosome. Thus, we find a pattern of X/A diversity and provided evidence implicating forces such as selection, genetic drift, migration, and admixture, which need to be invoked for explaining them in domestic species and their wild relatives.

Data Availability

The Illumina Ovine 50K SNP BeadChip used in analyses can be accessed in Kijas et al. (2012), Zhao et al. (2017), and the NextGen Project (http://projects.ensembl.org/nextgen/). The whole-genome sequence data are obtained from the NextGen Project and Yang et al. (2016).

Author Contributions

M.-H.L. designed and supervised the study. Z.-H.C., M.Z., and F.-H.L. analyzed the data with contribution from K.N. M.W.B. contributed to the data. M.-H.L., F.-H.L., M.Z., and Z.-H.C. wrote the paper with contribution from K.N. and M.W.B. All the authors reviewed and approved the final manuscript.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

This study was financially supported by the grants from the National Natural Science Foundation of China (Nos. 31272413 and 91731309) and the Taishan Scholars Program of Shandong Province (No. ts201511085).We acknowledge Rasmus Nielsen for helpful discussions and comments on an early version of this work. This study makes use of data generated by the NextGen Consortium. The European Union's Seventh Framework Programme (FP7/2010-2014) provided funding for the project under grant agreement no 244356—NextGen.

Literature Cited

Amaral AJ, et al. 2011. Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. PLoS One 6(4):e14782.

Andolfatto P. 2001. Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. Mol Biol Evol. 18(3):279–290.

Arbiza L, Gottipati S, Siepel A, Keinan A. 2014. Contrasting X-linked and autosomal diversity across 14 human populations. Am J Hum Genet. 94(6):827–844.

Barbato M, et al. 2017. Genomic signatures of adaptive introgression from European mouflon into domestic sheep. Sci Rep. 7(1):7623.

Begun DJ, Whitley P. 2000. Reduced X-linked nucleotide polymorphism in *Drosophila simulans*. Proc Natl Acad Sci U S A. 97(11):5960–5965.

Betancourt AJ, Kim Y, Orr HA. 2004. A pseudohitchhiking model of X vs. autosomal diversity. Genetics 168(4):2261–2269.

Burgos-Paz W, et al. 2013. Worldwide genetic relationships of pigs as inferred from X chromosome SNPs. Anim Genet. 44(2):130–138.

Carneiro M, et al. 2014. Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. Science 345(6200):1074–1079.

Casto AM, et al. 2010. Characterization of X-linked SNP genotypic variation in globally distributed human populations. Genome Biol. 11(1):R10.

Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex-chromosomes and autosomes. Am Nat. 130(1):113–146.

Charlesworth B. 2012. The role of background selection in shaping patterns of molecular evolution and variation: evidence from variability on the *Drosophila* X chromosome. Genetics 191(1):233–246.

Chen EY, et al. 2013. Enrichr: interactive and collaborative *HTML5* gene list enrichment analysis tool. BMC Bioinformatics 14:128.

Chessa B, et al. 2009. Revealing the history of sheep domestication using retrovirus integrations. Science 324(5926):532–536.

Corl A, Ellegren H. 2012. The genomic signature of sexual selection in the genetic diversity of the sex chromosomes and autosomes. Evolution 66(7):2138–2149.

Cotter DJ, Brotman SM, Wilson Sayres MA. 2016. Genetic diversity on the human X chromosome does not support a strict pseudoautosomal boundary. Genetics 203(1):485–492.

Cox MP, Karafet TM, Lansing JS, Sudoyo H, Hammer MF. 2010. Autosomal and X-linked single nucleotide polymorphisms reveal a steep Asian-Melanesian ancestry cline in eastern Indonesia and a sex bias in admixture rates. Proc R Soc B 277(1687):1589–1596.

Danecek P, et al. 2011. The variant call format and VCFtools. Bioinformatics 27(15):2156–2158.

Das PJ, Chowdhary BP, Raudsepp T. 2009. Characterization of the bovine pseudoautosomal region and comparison with sheep, goat, and other mammalian pseudoautosomal regions. Cytogenet Genome Res. 126(1–2):139–147.

Delaneau O, Howie B, Cox AJ, Zagury JF, Marchini J. 2013. Haplotype estimation using sequencing reads. Am J Hum Genet. 93(4):687–696.

Ellegren H. 2009. The different levels of genetic diversity in sex chromosomes and autosomes. Trends Genet. 25(6):278–284.

Emery LS, Felsenstein J, Akey JM. 2010. Estimators of the human effective sex ratio detect sex biases on different timescales. Am J Hum Genet. 87(6):848–856.

Esteve-Codina A, et al. 2011. Partial short-read sequencing of a highly inbred Iberian pig and genomics inference thereof. Heredity 107(3):256–264.

Evans BJ, Zeng K, Esselstyn JA, Charlesworth B, Melnick DJ. 2014. Reduced representation genome sequencing suggests low diversity on the sex chromosomes of Tonkean macaque monkeys. Mol Biol Evol. 31(9):2425–2440.
Galtier N. 2004. Recombination, GC-content and the human pseudoautosomal boundary paradox. Trends Genet. 20(8):347–349.
Goldberg A, Günther T, Rosenberg NA, Jakobsson M. 2017. Ancient X chromosomes reveal contrasting sex bias in Neolithic and Bronze Age Eurasian migrations. Proc Natl Acad Sci U S A. 114:2657–2662.
Gottipati S, Arbiza L, Siepel A, Clark AG, Keinan A. 2011. Analyses of X-linked and autosomal genetic variation in population-scale whole genome sequencing. Nat Genet. 43(8):741–743.
Hammer MF, et al. 2010. The ratio of human X chromosome to autosome diversity is positively correlated with genetic distance from genes. Nat Genet. 42(10):830–831.
Hedrick PW. 2007. Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. Evolution 61(12):2750–2771.
Heyer E, Segurel L. 2010. Looking for signatures of sex-specific demography and local adaptation on the X chromosome. Genome Biol. 11(1):203.
Ho SYW, et al. 2011. Time-dependent rates of molecular evolution. Mol Ecol. 20(15):3087–3101.
Huang DW, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37(1):1–13.
Huang DW, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 4(1):44–57.
Jensen P. 2014. Behavior genetics and the domestication of animals. Annu Rev Anim Biosci. 2:85–104.
Jiang Y, et al. 2014. The sheep genome illuminates biology of the rumen and lipid metabolism. Science 344(6188):1168–1173.
Jukes TH, Cantor C. 1969. Evolution of protein molecules. In: Munro HN, editor. Mammalian protein metabolism. New York: Academic. p. 21–132.
Keinan A, Mullikin JC, Patterson N, Reich D. 2009. Accelerated genetic drift on chromosome X during the human dispersal out of Africa. Nat Genet. 41(1):66–70.
Keinan A, Reich D. 2010. Can a sex-biased human demography account for the reduced effective population size of chromosome X in non-Africans? Mol Biol Evol. 27(10):2312–2321.
Kerje S, et al. 2004. The dominant white, dun and smoky color variants in chicken are associated with insertion/deletion polymorphisms in the *PMEL17* gene. Genetics 168(3):1507–1518.
Kijas JW, et al. 2012. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. PLoS Biol. 10(2):e1001258.
Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinformatics 15(1):356.
Kuleshov MV, et al. 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 44(W1):W90–W97.
Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 33(7):1870–1874.
Lamichhaney S, et al. 2012. Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. Proc Natl Acad Sci U S A. 109(47):19345–19350.
Lamichhaney S, et al. 2017. Parallel adaptive evolution of geographically distant herring populations on both sides of the North Atlantic Ocean. Proc Natl Acad Sci U S A. 114(17):E3452–E3461.
Larson G, Burger J. 2013. A population genetics view of animal domestication. Trends Genet 29(4):197–205.
Lasne C, Sgrò CM, Connallon T. 2017. The relative contributions of the X chromosome and autosomes to local adaptation. Genetics 205(3):1285–1304.
Li H, et al. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078–2079.
Librado P, et al. 2015. Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. Proc Natl Acad Sci U S A. 112(50):E6889–E6897.
Lu J, Wu CI. 2005. Weak selection revealed by the whole-genome comparison of the X chromosome and autosomes of human and chimpanzee. Proc Natl Acad Sci U S A. 102(11):4063–4067.
Lv FH, et al. 2014. Adaptations to climate-mediated selective pressures in sheep. Mol Biol Evol. 31(12):3324–3343.
Lv FH, et al. 2015. Mitogenomic meta-analysis identifies two phases of migration in the history of eastern eurasian sheep. Mol Biol Evol. 32(10):2515–2533.
McCue ME, et al. 2012. A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. PLoS Genet. 8(1):e1002451.
McVicker G, Gordon D, Davis C, Green P. 2009. Widespread genomic signatures of natural selection in hominid evolution. PLoS Genet. 5(5):e1000471.
Medugorac I, et al. 2017. Whole-genome analysis of introgressive hybridization and characterization of the bovine legacy of Mongolian yaks. Nat Genet. 49(3):470–475.
Nam K, et al. 2015. Extreme selective sweeps independently targeted the X chromosomes of the great apes. Proc Natl Acad Sci U S A. 112(20):6413–6418.
Nam K, et al. 2017. Evidence that the rate of strong selective sweeps increases with population size in the great apes. Proc Natl Acad Sci U S A. 114(7):1613–1618.
Narang P, Wilson Sayres MA. 2016. Variable autosomal and X divergence near and far from genes affects estimates of male mutation bias in Great Apes. Genome Biol Evol. 8(11):3393–3405.
Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A. 76(10):5269–5273.
Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. PLoS Genet. 2(12):e190.
Ramachandran S, Rosenberg NA, Feldman MW, Wakeley J. 2008. Population differentiation and migration: coalescence times in a two-sex island model for autosomal and X-linked loci. Theor Popul Biol. 74(4):291–301.
Rezaei HR, et al. 2010. Evolution and taxonomy of the wild species of the genus *Ovis* (Mammalia, Artiodactyla, Bovidae). Mol Phylogenet Evol. 54(2):315–326.
Rousset F. 2008. Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour. 8(1):103–106.
Ryder ML. 1983. Sheep and man. London: Duckworth.
Ryman N, Palm S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Mol Ecol Notes 6(3):600–602.
Sabeti PC, et al. 2007. Genome-wide detection and characterization of positive selection in human populations. Nature 449(7164):913–918.
Schaffner SF. 2004. The X chromosome in population genetics. Nat Rev Genet. 5(1):43–51.
Singh N, Petrov DA. 2007. Evolution of gene function on the X chromosome versus the autosomes. Genome Dyn. 3:101–118.
Szpiech ZA, Hernandez RD. 2014. Selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Mol Biol Evol. 31(10):2824–2827.
Van Laere AS, Coppieters W, Georges M. 2008. Characterization of the bovine pseudoautosomal boundary: documenting the evolutionary

history of mammalian sex chromosomes. Genome Res. 18(12):1884–1895.
VanBuren R, et al. 2016. Extremely low nucleotide diversity in the X-linked region of papaya caused by a strong selective sweep. Genome Biol. 17(1):230.
Veeramah KR, Gutenkunst RN, Woerner AE, Watkins JC, Hammer MF. 2014. Evidence for increased levels of positive and negative selection on the X chromosome versus autosomes in humans. Mol Biol Evol. 31(9):2267–2282.
Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet. 7(8):645–653.
Voight BF, Kudaravalli S, Wen XQ, Pritchard JK. 2006. A map of recent positive selection in the human genome. PLoS Biol. 4(3):e72.
Waldman YY, et al. 2016. The genetics of Bene Israel from India reveals both substantial Jewish and Indian ancestry. PLoS One 11(3):e0152056.
Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38(16):e164.
Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38(6):1358–1370.
White MA, Ikeda A, Payseur BA. 2012. A pronounced evolutionary shift of the pseudoautosomal region boundary in house mice. Mamm Genome 23(7–8):454–466.
Xu SS, Li MH. 2017. Recent advances in understanding genetic variants associated with economically important traits in sheep (*Ovis aries*) revealed by high-throughput screening technologies. Front Agr Sci Eng. 4(3):279–288.
Yang J, et al. 2016. Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. Mol Biol Evol. 33(10):2576–2592.
Zeder MA. 2008. Domestication and early agriculture in the Mediterranean Basin: origins, diffusion, and impact. Proc Natl Acad Sci U S A. 105(33):11597–11604.
Zhao YX, et al. 2017. Genomic reconstruction of the history of native sheep reveals the peopling patterns of nomads and the expansion of early pastoralism in East Asia. Mol Biol Evol. 34(9):2380–2395.

Associate editor: Ya-Ping Zhang