

SCIENTIFIC REPORTS



OPEN

Global genomic diversity and conservation priorities for domestic animals are associated with the economies of their regions of origin

Min Zhang^{1,2}, Wei-Feng Peng^{2,3}, Xiao-Ju Hu^{2,3}, Yong-Xin Zhao^{2,3}, Feng-Hua Lv² & Ji Yang²

Domestic animals play a key role in human survival and the development of civilization. However, the genetic resources of domestic animals are facing an alarming rate of erosion due to socioeconomic changes, economic globalization and financial constraints. In this study, through genome-wide SNP analysis, we estimated the heterozygosity, inbreeding coefficient, effective population size, and runs of homozygosity to identify the breeds facing the risk of extinction for sheep and cattle across the world. In particular, we quantified the contribution of 97 sheep breeds and 53 cattle breeds to genomic diversity (within-breed, between-breed and total) and prioritized the breeds for conservation. Additionally, we compared the average values of genomic diversity between breeds from regions (or countries) in different economic categories (underdeveloped, developing and developed), and found that breeds in developed regions exhibit significantly higher levels of total genomic diversity than those in underdeveloped and developing regions. Altogether, our results suggested that conservation priority should be given to breeds in developed regions to secure the future genomic diversity hotspots of domestic animal resources.

Farm animals contribute in many ways to human survival and economic development, particularly as societies evolve from subsistence agriculture into market-based economies¹. However, in the past two decades, farm animal genetic resources have faced an alarming rate of extinction and genetic erosion, as approximately 200 uniquely adapted breeds have become extinct and up to 30% of global livestock breeds (approximately 1,200–1,500) are currently endangered². This loss of diversity is attributed to multiple causes, including regional economic forces such as the decline in the economic viability of traditional livestock production systems³, destruction of the native habitats (e.g., grazing lands) of livestock breeds⁴, rapid dissemination of uniform and highly productive breeds at the expense of native genetic resources⁵ and the extension of markets and economic globalization⁶. In this context, an integrated analysis of the genomic diversity of farm animals and the associated economies can aid an understanding of economic impact on the changes in farm animal genomic diversity in recent years. Additionally, it will contribute to the development of governmental interventions and the implementation of conservation priorities that are necessary to ensure that ongoing agricultural development will be compatible with the conservation and sustainable use of local livestock diversity⁷.

Here, we ranked the conservation priorities of 97 native sheep breeds and 53 native cattle breeds from around the world based on genome-wide single-nucleotide polymorphism (SNP) data for the first time. Firstly, we calculated heterozygosity, inbreeding coefficient, and runs of homozygosity (ROH) for each sheep and cattle breed to detect the small and inbred populations that are facing the risk of extinction. Moreover, we implemented the first association analysis between economic indexes and genomic diversity by studying a total of 2,941 animals from 150 native breeds of two globally distributed livestock species: sheep (1,910 animals, 97 breeds; Supplementary Table 1) and cattle (1,031 animals, 53 breeds; Supplementary Table 2). Specifically, the estimations of genomic diversity contributions for sheep and cattle were based on SNP array data genotyped from the Ovine SNP50 (or Infinium HD) and Bovine SNP50 BeadChips, respectively. Through the molecular coancestry⁸ and allelic richness

¹School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, 230027, China. ²CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences (CAS), Beijing, 100101, China. ³University of Chinese Academy of Sciences (UCAS), Beijing, 100049, China. Correspondence and requests for materials should be addressed to J.Y. (email: yangji@ioz.ac.cn)

Parameters	Categories	Number of breeds	
		Sheep	Cattle
N_e	$N_e \leq 100$	75	33
	$100 < N_e < 200$	10	10
	$N_e \geq 200$	12	10
F_{ROH}	$F_{ROH} \geq 0.1$	10	3
	$0.05 < F_{ROH} < 0.1$	24	14
	$F_{ROH} \leq 0.05$	63	36
N_{ROH}	$N_{ROH} \geq 200$	31	12
	$150 < N_{ROH} < 200$	11	13
	$N_{ROH} \leq 150$	55	28
L_{ROH}	$L_{ROH} \geq 20$	6	1
	$15 < L_{ROH} < 20$	37	6
	$L_{ROH} \leq 15$	54	46

Table 1. The number of sheep and cattle breeds in different categories of population genetic parameters.

approaches⁹, we calculated the breed contribution to genomic diversity at three levels: within-population diversity (ΔGD_{WS}), between-population diversity (ΔGD_{BS}), and total genomic diversity (ΔGD_T) for each breed after its removal (Supplementary Tables 1 and 2). For region-wide breeds, the estimation of genomic diversity contributions was implemented by excluding all of the breeds within the region from the total breeds (Supplementary Tables 3 and 4). We then compared the ΔGD_{WS} , ΔGD_{BS} and ΔGD_T among the regions of origin (or countries) in terms of the annual gross domestic product (GDP) per capita (GDPPC). We also took into account the increase in the GDP per capita ($\Delta GDPPC$) for the regions during the past two decades in a dual-regression analysis to correlate the contribution of the genomic diversity of the breeds and the past economic development in a region. To facilitate comparisons among the regions with different economies, we grouped the regions into three categories (i.e., underdeveloped, developing and developed) according to the average GDPPC (A-GDPPC) during the years 1993–2013. Based on all the above analyses, we were finally able to determine the contribution of each breed to the maximum amount of genomic diversity and establish the core sets of breeds with higher priority for conservation.

Results

Genetic diversity pattern and inbreeding coefficient. Heterozygosity is a measure of genetic variation within a population. We estimated the observed (H_O) and expected (H_E) heterozygosity for each sheep and cattle breed. For sheep breeds, the highest H_O value was observed in the RAA (0.394) breed from Spain, and the lowest was in the SPS breed (0.252) from Yunnan province of China. The highest and lowest expected (H_E) heterozygosity were also observed in the RAA breed (0.385) and SPS breeds (0.258), respectively (Supplementary Table 1). For the cattle breeds, CANC breed from Brazil and PRP breed from France had the highest H_O (0.362) and H_E (0.343) values, respectively, and BALI breed from Indonesia presented the lowest H_O (0.061) and H_E (0.063) values (Supplementary Table 2). For the estimation of inbreeding coefficients (F_{IS}) of sheep breeds, BGE breed had the highest F_{IS} value (0.157) and DSH breed had the lowest F_{IS} value (-0.065) (Supplementary Table 1). In cattle breeds, the SOM (0.050) and ND2 (-0.108) presented the highest and lowest F_{IS} values, respectively (Supplementary Table 2).

Effective population size. The results for effective population size (N_e) for sheep and cattle were calculated based on the linkage disequilibrium method and presented in the Supplementary Tables 1 and 2. In sheep, N_e ranged from 3,074 (BYK) down to 3 (WNS), while in cattle the value ranged from 1,429 (ZBO) down to 1 (BALI). The high degree of inbreeding may be the reason for the extremely low N_e value of BALI, as evidenced by the high F_{IS} and F_{ROH} value observed in this breed (Supplementary Table 2). To glean further details, we divided the N_e values into three categories: $N_e \leq 100$, $100 < N_e < 200$, and $N_e \geq 200$, and the number of breeds within each of the three categories were summarized in Table 1 (also see Figs 1a and 2a, Supplementary Figs 1a and 2a and Supplementary Tables 1 and 2).

Runs of homozygosity (ROH). In this study, we calculated the ROH statistics for each breed of sheep and cattle, and subsequently the inbreeding coefficient based on ROH (F_{ROH}) was estimated. The F_{ROH} values ranged from 0.0004 (WZS) to 0.2016 (WIL) in sheep breeds (Supplementary Table 1) and from 0.0001 (ZEB) to 0.6842 (BALI) in cattle breeds (Supplementary Table 2). Regarding the number of ROH (N_{ROH}) above 500 kb, the sheep breed of WIL ($N_{ROH} = 780$) showed the highest N_{ROH} value and WZS ($N_{ROH} = 2$) had the lowest N_{ROH} value (Supplementary Table 1). For the cattle breeds, ZEB ($N_{ROH} = 1$) and BALI ($N_{ROH} = 1,127$) breeds showed the lowest and highest N_{ROH} value, respectively (Supplementary Table 2). Finally, we also calculated the average length of ROH above 500 kb for each breed (L_{ROH}). For sheep and cattle breeds, L_{ROH} values ranged from 26.9 Mb (HLS) to 9.0 Mb (SOA), and 21.8 Mb (BOR) to 5.1 Mb (ZEB), respectively (Supplementary Tables 1 and 2). To obtain more details, F_{ROH} , N_{ROH} and L_{ROH} values were also divided into three categories: F_{ROH} categories ($F_{ROH} \geq 0.10$, $0.05 < F_{ROH} < 0.10$, and $F_{ROH} \leq 0.05$), N_{ROH} categories ($N_{ROH} \geq 200$, $150 < N_{ROH} < 200$, and $N_{ROH} \leq 150$) and L_{ROH} categories ($L_{ROH} \geq 20$ Mb, $15 \text{ Mb} < L_{ROH} < 20$ Mb, and $L_{ROH} \leq 15$ Mb), and the number of breeds belong to each

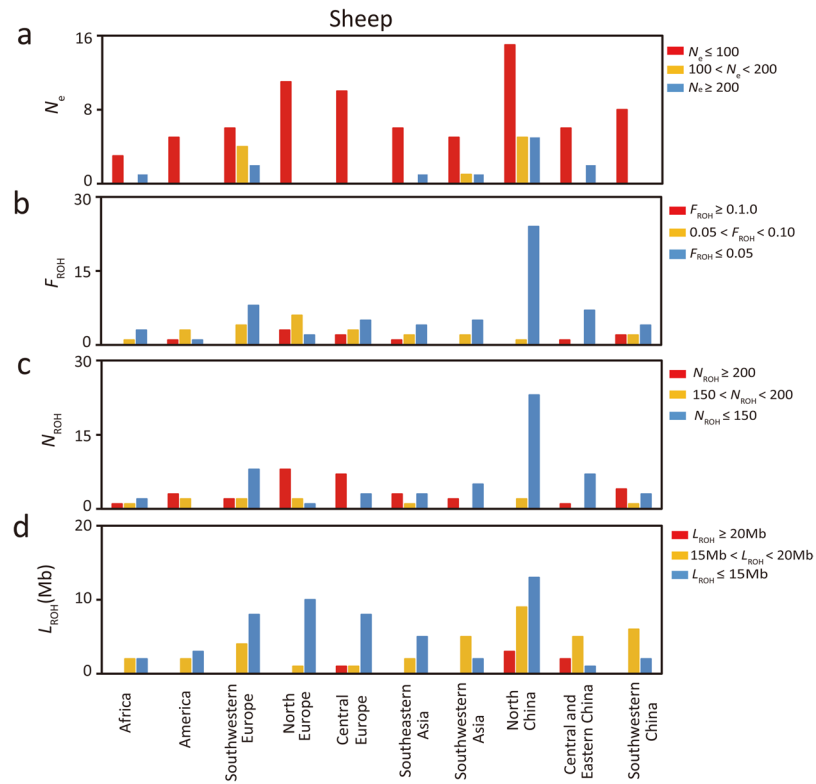


Figure 1. Distribution patterns of N_e (a) N_{ROH} (b) L_{ROH} (c) F_{ROH} (d) found in 97 sheep breeds. N_e : effective population size. F_{ROH} : inbreeding coefficient based on runs of homozygosity. N_{ROH} : number of ROH. L_{ROH} : mean length of ROH. The Y axis represents the number of breeds that fall into three different ranges of the statistics.

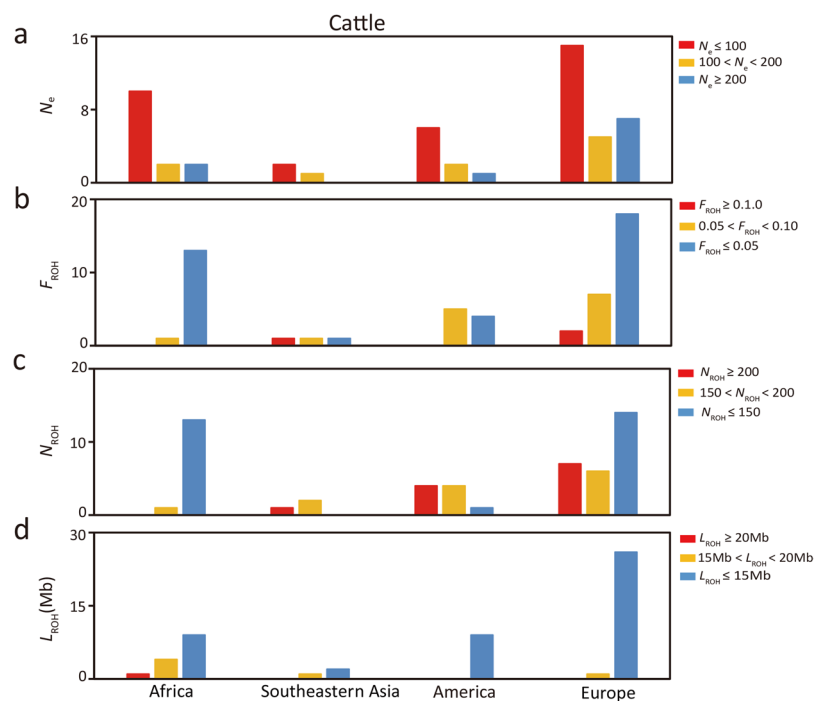


Figure 2. Distribution patterns of N_e (a) N_{ROH} (b) L_{ROH} (c) F_{ROH} (d) found in 53 cattle breeds. N_e : effective population size. F_{ROH} : inbreeding coefficient based on runs of homozygosity. N_{ROH} : number of ROH. L_{ROH} : mean length of ROH. The Y axis represents the number of breeds that fall into three different ranges of the statistics.

of the categories were shown in Table 1 (also see Figs 1 and 2, Supplementary Figs 1 and 2 and Supplementary Tables 1 and 2).

Genetic structure. The relationship between breeds was investigated using the NJ tree based on the Reynolds distance estimated from allele frequency differences, and branch lengths were highly variable. The 97 sheep breeds were clearly divided into African, Asian (Southwestern Asian, Southeastern Asian, and China), and European sheep breeds (Supplementary Fig. 3a). However, sheep from the Americas were clustered with European-derived sheep breeds (Supplementary Fig. 3a). The 53 cattle breeds were clearly divided into European and African cattle breeds, although the American and Southeastern Asian cattle breeds exhibited an admixture with cattle breeds from Europe and Africa (Supplementary Fig. 3b).

Conservation priorities. The mean values obtained using the molecular coancestry and allelic richness methods showed ranges for $\Delta\text{GD}_{\text{WS}}$, $\Delta\text{GD}_{\text{BS}}$, and $\Delta\text{GD}_{\text{T}}$ in sheep from -0.200 (SPS) to 0.077 (RAA), -0.067 (KAZ) to 0.218 (SOA), and -0.044 (SPS) to 0.042 (DSH), respectively (Supplementary Table 1). In cattle, these values ranged from -0.781 (BALI) to 0.257 (CANC), -0.198 (CANC) to 0.670 (BALI) and -0.176 (LAG) to 0.172 (HFD), respectively (Supplementary Table 2).

As representative of the main cattle breeds reared around the world, six cattle breeds (HFD, HO, AN, ANR, NRC, and BEFM; 11.3%) showed a $\Delta\text{GD}_{\text{T}} > 0.1$ and 11 (20.8%) breeds showed a $\Delta\text{GD}_{\text{T}} < -0.1$ in this study (Supplementary Table 2 and Supplementary Fig. 4f). However, none of the native sheep breeds presented extreme values of $\Delta\text{GD}_{\text{T}} > 0.1$ or < -0.1 (Supplementary Table 1 and Supplementary Fig. 4c). This observation implied a more diverse genetic constitution of sheep compared with cattle. Another possible reason is the differences in number of sheep and cattle breeds analyzed, as the contribution of a single breed to genomic diversity is expected to be lower in a larger sample.

Comparison of conservation priorities in three different economic categories. According to the average GDP per capita (A-GDPPC) during 1993 to 2013, we divided countries or regions into three economic categories: underdeveloped countries/regions, developing countries/regions, and developed countries/regions. For sheep, English breeds showed the highest $\Delta\text{GD}_{\text{T}}$ (0.259) value, Xinjiang breeds showed the highest $\Delta\text{GD}_{\text{WS}}$ (0.278) value, and South African breeds were observed with the highest $\Delta\text{GD}_{\text{BS}}$ (0.179) value. In contrast, Bangladeshi breeds showed the lowest $\Delta\text{GD}_{\text{WS}}$ (-0.205) value, and Xinjiang breeds had the lowest $\Delta\text{GD}_{\text{BS}}$ (-0.445) and $\Delta\text{GD}_{\text{T}}$ (-0.167) values (Supplementary Table 3). For cattle, $\Delta\text{GD}_{\text{WS}}$ ranged from 2.271 (France) to -1.421 (Indonesia), $\Delta\text{GD}_{\text{BS}}$ ranged 1.134 (Indonesia) to -1.225 (France), and $\Delta\text{GD}_{\text{T}}$ ranged from 1.047 (France) to -0.293 (Pakistan) (Supplementary Table 4). Further comparisons revealed significant differences in $\Delta\text{GD}_{\text{WS}}$ and $\Delta\text{GD}_{\text{T}}$ between breeds from the three economic categories. For instance, both species showed that the breeds from developed regions had the highest average value of $\Delta\text{GD}_{\text{T}}$, followed by those from developing regions and underdeveloped regions (Fig. 3). In addition, the breeds in developing and developed regions exhibited significantly ($P < 0.05$) higher $\Delta\text{GD}_{\text{WS}}$ values than those in underdeveloped regions, whereas the $\Delta\text{GD}_{\text{BS}}$ estimates for breeds from different regions within the three economic categories were not significantly different from each other (Fig. 3a). We also examined the impact of economic growth, which is typically measured as the rate of change in the average annual GDP per capita (ΔGDPPC), on the gain or loss of genomic diversity. Our results showed that $\Delta\text{GD}_{\text{WS}}$ and $\Delta\text{GD}_{\text{T}}$ presented significant associations with ΔGDPPC during the past two decades (sheep: $R^2 = 0.087$, $P = 0.073$ ($\Delta\text{GD}_{\text{WS}}$), $R^2 = 0.278$, $P < 0.01$ ($\Delta\text{GD}_{\text{T}}$); cattle: $R^2 = 0.235$, $P < 0.05$ ($\Delta\text{GD}_{\text{WS}}$); $R^2 = 0.318$, $P < 0.01$ ($\Delta\text{GD}_{\text{T}}$); Fig. 4a,c,d,f). In both species, $\Delta\text{GD}_{\text{WS}}$ and $\Delta\text{GD}_{\text{T}}$ displayed a rapid increase for $\Delta\text{GDPPC} < 20,000$ and a slow increase thereafter (Fig. 4).

Discussion

Heterozygosity is an important factor for estimating the genetic variability in domestic animals. Heterozygosity represents the genetic potential and adaptation abilities to the natural environment¹⁰. In this study, we found that the livestock in developing and developed regions have higher heterozygosity than those in underdeveloped regions or some island populations. For example, sheep breeds from Southwestern Europe and Northern China obviously harbor higher observed (H_{O}) and expected (H_{E}) heterozygosity than those from Africa and southwestern China, which belong to underdeveloped regions (Supplementary Table 1). Similarly, cattle breeds from Europe and America possess higher H_{O} and H_{E} values than those from Africa and Asia (Supplementary Table 2). Besides economic factors, another reasonable explanation is that breeds with high heterozygosity are from regions near the domestication center or trade routes^{11–15}. Notably, we found that the island breeds always show low heterozygosity. For example, BALI breed, which is distributed on the Bali island of Indonesia, exhibits the lowest heterozygosity ($H_{\text{O}} = 0.061$ and $H_{\text{E}} = 0.063$). Also, GNS ($H_{\text{O}} = 0.299$ and $H_{\text{E}} = 0.297$) and JER ($H_{\text{O}} = 0.290$ and $H_{\text{E}} = 0.280$) breeds from Guernsey and Jersey islands near England harbor lower heterozygosity than breeds from the European continent (Supplementary Table 2). These results suggested that these island breeds have suffered population decline and their genetic variation has been reduced.

Effective population size (N_{e}) has been regarded as a popular parameter to investigate small populations and endangered species. According to the 50/500 rule of thumb¹⁶, N_{e} values under 50 means that the population is likely to face a serious genetic threat after 5 or more generations. This is because small populations cannot easily avoid inbreeding depression. N_{e} values above 500 means that the population can maintain an evolutionary potential for longer periods of time. However, Frankham *et al.*¹⁷ have proposed a new criterion that N_{e} values under 100 and above 1,000 are suitable to control inbreeding and maintain high genetic diversity, respectively¹⁷. Based on the new criterion, more than half of the breeds for both species in this study are facing the risk of extinction as their N_{e} values were under 100 (Figs 1a and 2a, Supplementary Figs 1a and 2a and Supplementary Tables 1 and 2). The probable cause of this observation was human intervention in the process of livestock development, such as

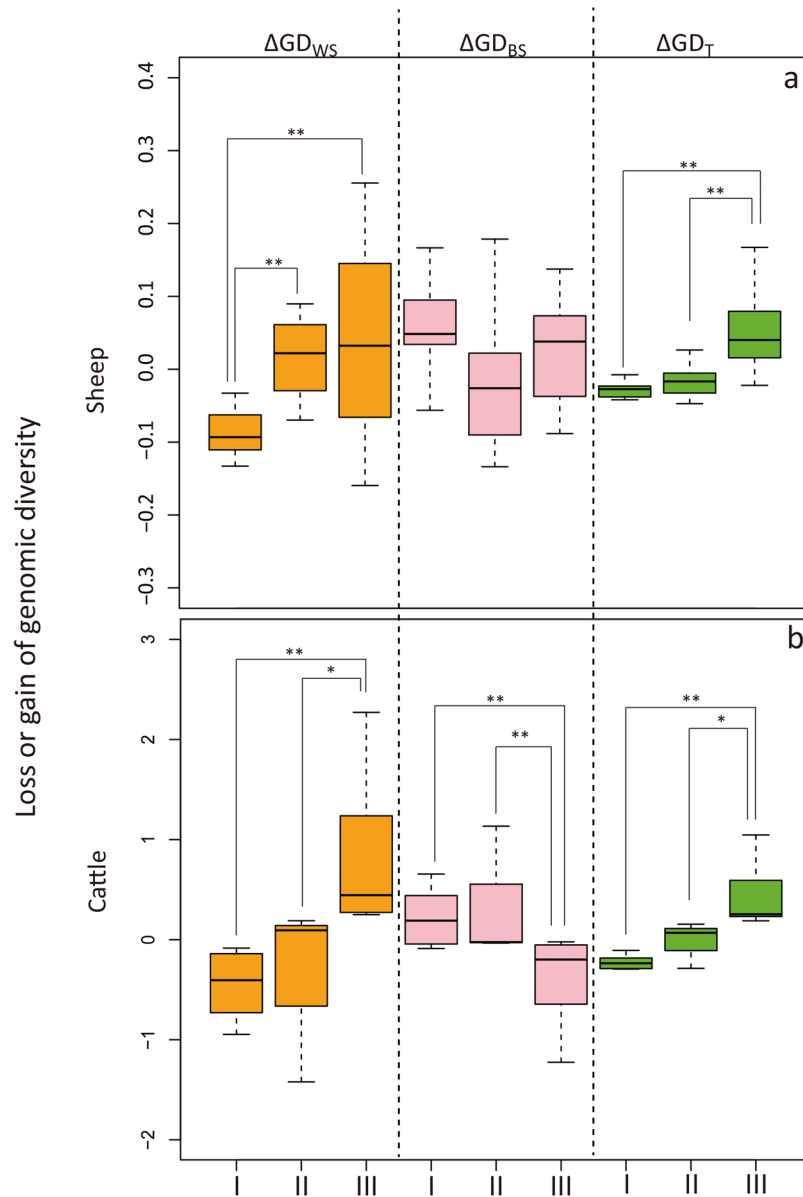


Figure 3. Comparisons of the genomic diversity contributions of sheep (a) and cattle (b) breeds between regions (or countries) within the three economic categories. Three economic categories: I, underdeveloped; II, developing; III, developed. Three genomic diversity contributions: within-population, ΔGD_{WS} (yellow); between-population, ΔGD_{BS} (pink); total genomic diversity, ΔGD_T (green). ** $P < 0.01$; * $P < 0.05$. The loss or gain of genomic diversity is indicated by the average values of estimates obtained using the molecular coancestry and allelic richness methods.

introduction of breeds and intensive breeding^{18–22}. Overall, we can see that most sheep and cattle breeds do not show a long-term stable status, and thus, we need to pay more attention and carry out appropriate conservation strategies on these native breeds with small effective population size.

ROH is a key parameter to detect the level of matings between related individuals in a closed or small population, and the number and length of ROH could indicate the extent of recent inbreeding and artificial selection^{23,24}. Because of its utility, ROH has been used to investigate genome-wide autozygosity regions associated with some important traits and diseases in domestic animals and humans^{25,26}. Inbreeding is among the major causes that lead to an increase in homozygosity, which in turn, reduce fitness and accumulate more recessive alleles that may lead to expression of unfavorable phenotypes^{27,28}. According to the inbreeding coefficient based on runs of homozygosity (F_{ROH}), we have identified 10 sheep breeds (WIL, BOR, WNS, BCS, ZTS, BGE, DSH, BHM, EFB and YXZ) and three cattle breeds (BALI, SH and JER) with an F_{ROH} value above 0.10. In addition, these breeds also showed a large number of runs of homozygosity ($N_{ROH} \geq 200$) (Supplementary Tables 1 and 2). Together these results suggest that these breeds were founded as small populations or the populations recently experienced severe inbreeding. Previous studies have provided evidence that domestic animals have suffered serious

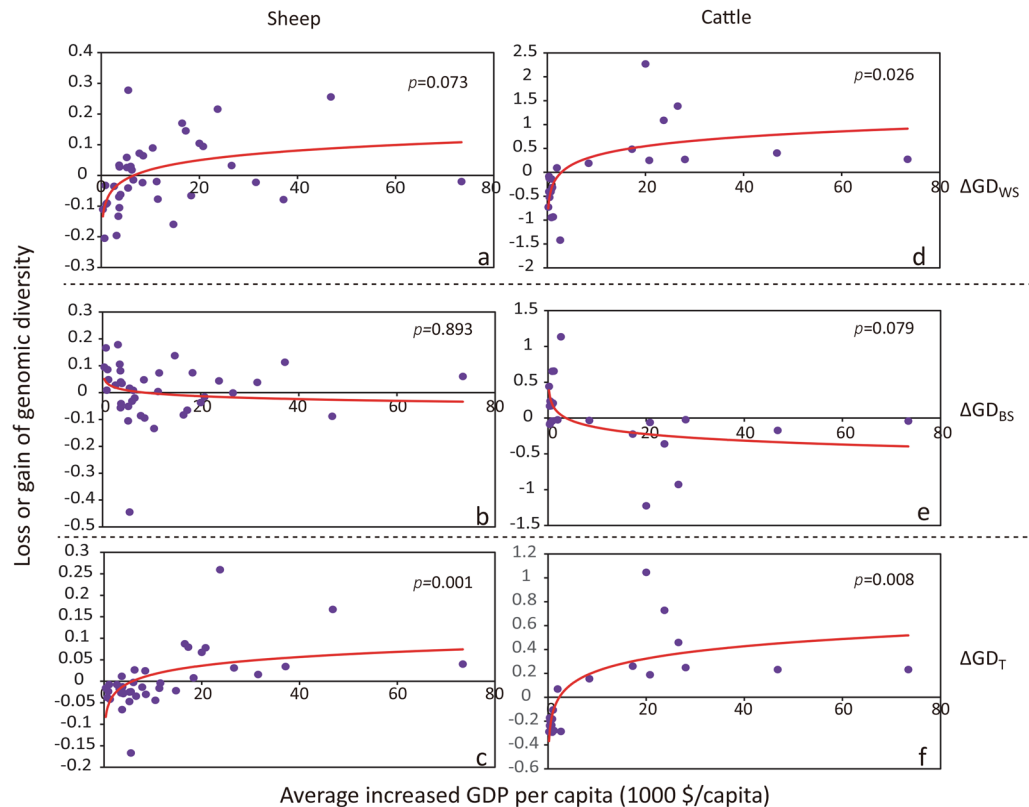


Figure 4. Regression between the genomic diversity contributions and the increase in GDP per capita (ΔGDPPC) during the years 1993–2013. (a–c) Within-population ($\Delta\text{GD}_{\text{WS}}$), $R^2 = 0.087$, $P > 0.05$; between-population ($\Delta\text{GD}_{\text{BS}}$), $R^2 = 0.001$, $P > 0.05$; total genomic diversity ($\Delta\text{GD}_{\text{T}}$), $R^2 = 0.278$, $P < 0.01$ in sheep. (d–f) Within-population ($\Delta\text{GD}_{\text{WS}}$), $R^2 = 0.0235$, $P < 0.05$; between subpopulation ($\Delta\text{GD}_{\text{BS}}$), $R^2 = 0.153$, $P > 0.05$; total genetic diversity ($\Delta\text{GD}_{\text{T}}$), $R^2 = 0.318$, $P < 0.01$ in cattle. The loss or gain of genomic diversity is indicated by the average values of the estimates obtained using the molecular coancestry and allelic richness methods.

inbreeding and the effective population size is declining under the pressure of artificial selection^{29,30}. In our study, most of the sheep and cattle breeds had a ROH length above 10 Mb, suggesting they share recent ancestors 3 generations ago^{31,32}. Also, there was a negative correlation between F_{ROH} and N_e , especially in the sheep breeds. Sheep breeds from North China had a lower F_{ROH} and a higher N_e than other regions, while in cattle breeds the lowest F_{ROH} and the highest N_e appeared in Africa and Europe. Interestingly, although some sheep and cattle breeds from the developed regions show high F_{ROH} , N_{ROH} and L_{ROH} and low N_e values, they make contributions to genetic diversity. For example, 6 sheep breeds (BOR, DSH, ISF, SOA, WIL and GAL) have a N_e value under 200, but they exhibit high $\Delta\text{GD}_{\text{BS}}$ and positive $\Delta\text{GD}_{\text{T}}$ values (Supplementary Table 1). Also, three cattle breeds (SH, GNS and JER) present the same pattern as the sheep breeds. Therefore, these breeds should be subjected to conservation measures to avoid the erosion of their genetic diversity. In addition, the breeds from developing and underdeveloped regions with low N_e values, such as sheep breeds BGE, GAR, ZTS and SPS, and cattle breeds LAG, ND2, BOR, ZMA and BALI (Supplementary Tables 1 and 2), are strongly differentiated from the continental breeds (Supplementary Fig. 3), implying that they may have some special economic traits to be used in breeding programs.

Population structure is a useful information for humans to design effective strategies to improve the conservation of farm animal genetic resources. In this study, we performed a comprehensive analysis of population structure of sheep and cattle breeds worldwide. The 97 sheep breeds can be clearly divided into six geographic regions (Europe, Southwestern Asia, Southeastern Asia, China, America and Africa) (Supplementary Fig. 3a). Previous research on 74 diverse sheep breeds collected globally documented a high level of admixture within the European sheep, which showed high co-ancestry with most other breeds³³. Compared to the previous study, our data set incorporated sheep breeds from China, and identified the breeds with large distinctness among the sheep breeds. Based on the results of the NJ tree, the breeds from Southeastern Asia, Africa, Southwestern China and several breeds (WIL, BRL, DSH, EFB, ISF and GAL) from Europe have longer branches (Supplementary Fig. 3a), suggesting these breeds are isolated breeds and possess unique genetic characteristics. The investigation of genetic structure of cattle reveals that cattle have two domestication centers: Fertile Crescent and Indus valley¹⁵. The African taurine has large divergence from European taurine and Asian indicine because of their large portion of wild African auroch ancestry³⁴. This is consistent with our results that African cattle show apparent divergence from European cattle (Supplementary Fig. 3b). In addition, 7 African (LAG, BAO, ND2, ZMA, NDAM, SOM and ZEB) and 3 Southeastern Asian breeds (SAHW, GIR and BALI) make a large contribution to the between-breed

genetic diversity (ΔGD_{BS}) (Supplementary Table 2). These results reflect the unique genetic resources carried by these breeds.

From the results of loss or gain of genetic diversity, coupled with the evidence of the above statistics (i.e., heterozygosity, N_e , ROH and population structure), our most prominent finding was that breeds from developed regions exhibit higher total genomic diversity than those in underdeveloped and developing regions, and thus, should be given a higher conservation priority (Fig. 3). Since the sample of this study does not cover the global distributions of sheep and cattle, future studies based on more extensive and representative samples are needed to verify this finding. For the potential factors affecting conservation priority, whether the economic situation measured based on the annual GDP per capita plays a direct role in livestock conservation is unclear. One hypothesis is that a better economy induces more effective conservation measures, including higher subsidies for raising local livestock breeds and more financial investment in conservation programs. For example, the European Union has long recognized the importance of conserving animal genetic resources (AnGR) and, in 1992, initiated a policy of economic incentives for farmers keeping native breeds, under EC regulations 2078/92 and 1257/99³⁵. However, breeds from underdeveloped regions, such as African countries (EMZ, NQA, RMA and RDA) and the Yunnan-Kweichow Plateau of China (ZTS, WGS, SPS, TCS, DQS, NLS and WNS), where little or no financial support has been provided for local livestock conservation measures, showed extreme negative values of within-breed genomic diversity (Supplementary Table 1 and Supplementary Fig. 4a). This observation could be due to factors such as long-term genetic isolation, genetic drift, or inbreeding, and thus, effective conservation measures preventing genetic drift in small-sized and isolated breeds and promoting reasonable gene flow with outside breeds should be implemented for the native breeds in underdeveloped regions. A second hypothesis is that economic growth is one causative factor for the change in the livestock production system, possibly through effects on breeding programs. Rapid economic growth could have led to an increased need for more productive and synthetic livestock breeds³⁶ and therefore contributed to the increased ΔGD_T and ΔGD_{WS} and decreased ΔGD_{BS} values observed in livestock (Fig. 4). Native breeds in developed regions exhibited high values of both ΔGD_{WS} and ΔGD_T (Fig. 3), which could also be due to the improvement of breeding programs by avoiding indiscriminate crossbreeding³⁷. However, breeds in developing regions such as Xinjiang of China showed high values of ΔGD_{WS} but negative values of ΔGD_T (Supplementary Table 1 and Supplementary Fig. 4a,c), which could be explained by the development scheme of extensive and intensive crossbreeding of sheep with exotic breeds adopted in past decades³⁸. Apart from the potential impact of economic situation on livestock conservation, several other factors may also account for the conservation status of domestic animals. For instance, the development of the breed concept, which originated 200–250 years ago in Europe, has probably affected the number of breeds differentiated and selected within the region. Also, transboundary exchanges of breeds may have played a fundamental role in shaping the genetic diversity of breeds, especially in American and Australian countries.

Another important finding regarding genetic diversity was that breeds in developing regions have been facing the erosion of between-breed genomic diversity (ΔGD_{BS}). Both species are transboundary animals, and many highly productive commercial breeds have been transferred worldwide to upgrade local breeds^{39–42}. With the rapid economic development and increase in the human population in developing regions, the share of trade in genetic material from developed to developing countries increased from 20% in 1995 to 30% in 2005^{43,44}. Long-term intensive crossbreeding programs using imported exotic breeds have narrowed the diversity of local genetic materials and led to the loss of between-breed genomic diversity in developing regions⁴⁵. In northern China, the estimates of both ΔGD_{BS} and ΔGD_T were negative, suggesting that northern Chinese sheep have suffered severe genetic erosion induced by exotic breeds in the past, which has resulted in the loss of both total and within-breed genomic diversity (Supplementary Fig. 4b,c). Similar scenarios were also observed in the breeds of Iran, Turkey and Pakistan (Supplementary Fig. 4a,c). However, livestock in underdeveloped regions (including remote and isolated island areas) are commonly used to service the local community^{46,47}, being subjected to little or no influence from exotic breeds, and these populations therefore showed high ΔGD_{BS} values (Supplementary Fig. 4b,e). For example, the diverse and isolated topography of the Yunnan-Kweichow Plateau of China has contributed to the genetically isolated populations of local sheep breeds, which show high and positive ΔGD_{BS} values, but negative ΔGD_T values (Supplementary Fig. 4b,c). Similar results were observed in cattle from mainland Africa and Madagascar (Supplementary Fig. 4e,f). In these scenarios, breeds with high levels of within-breed genomic diversity (ΔGD_{WS}) should be ranked with high conservation priorities to preserve closed populations that will be capable of coping with future challenging environments or diversified food production. However, breeds with large divergence should also give protection measures to prevent their extinction due to their unique genetic resources and special adaptation to local environment.

As the ovine and bovine SNP BeadChips have generally been developed on European breeds, it is expected that ascertainment bias would impact the estimates of genetic diversity metrics. However, our results should not be significantly affected by ascertainment bias for the following reasons. On the one hand, Herráez *et al.*⁴⁸ have demonstrated that the effect of ascertainment bias could be countered if the SNPs with high LD were removed. Based on this finding, we recalculated the heterozygosity and ΔGD_T (loss or gain of genetic diversity) by using the LD-pruned data, and obtained similar results with those not controlling for LD (Supplementary Fig. 5). On the other hand, Kijas *et al.*³³ calculated genetic diversity of sheep by using two types of SNP chip data which were genotyped by Roche 454 and Illumina GA sequencing, respectively. The Roche 454 SNPs were developed primarily using animals of European origin, while the Illumina GA SNPs were designed using a larger number of animals selected from multiple regions. The results showed that ascertainment bias was unlikely to heavily influence genetic diversity between breeds and regions. From all the evidence presented above, we believe that ascertainment bias should have little influence on our results.

In conclusion, we show that native sheep and cattle breeds from developed regions in this study make a greater contribution to total genomic diversity than those from developing and underdeveloped regions; thus, a higher conservation priority should be given to these breeds. In contrast, most of the breeds in underdeveloped and

developing regions make a small or no contribution at all to total genomic diversity and therefore receive relatively lower rankings in terms of conservation priority. Nevertheless, we noticed that breeds from underdeveloped countries tend to contribute more to between-breed diversity (Fig. 3). This is also a conservation objective, as many of these small, isolated and peripheral populations may have valuable variation (e.g., private alleles) that should be conserved. Also, it should be noted that the lack of samples in some distribution areas of sheep and cattle limit the generality of the results. The mechanisms underlying these observations are not well understood, but the regional annual GDP per capita and its growth rate, which are two important indicators of economic development, could be the missing link and help to explain the differences in the genomic diversity contributions of livestock breeds from different regions. In addition to genomic diversity, other factors, such as sociocultural value, local adaptation and the special uses of some livestock breeds, should also be considered in the development of conservation programs. With an awareness that the conservation recommendations made from the analysis of the contribution of breeds to diversity depend on the conservation interests and scopes, our results could help local authorities and animal breeders to ensure the effective and sustainable utilization of animal genetic resources in the future.

Methods

The methods were carried out in accordance with the approved guidelines of the Good Experimental Practices adopted by the Institute of Zoology, Chinese Academy of Sciences. All experimental procedures and animal collections were conducted under a permit (No. IOZ13015) approved by the Committee for Animal Experiments of the Institute of Zoology, Chinese Academy of Sciences, China.

Ovine sample collection and DNA extraction. Blood or tissue samples from 783 animals representing 40 native Chinese sheep breeds were collected in this study (Supplementary Table 1). In all cases, particular efforts were made, based on both pedigree information and the knowledge of local herdsmen, to ensure that the animals were unrelated and typical of their breeds. Whole blood samples were collected into evacuated tubes containing EDTA and ear marginal tissues were collected and stored in 2-ml microcentrifuge tubes containing 75% ethanol. Genomic DNA was extracted from the tissue samples using the standard phenol/chloroform protocol⁴⁹ or from blood using the General AllGen Kit (Tiangen Biotech, Beijing, China) following the manufacturer's instructions.

Ovine SNP datasets genotyping. Of the 40 Chinese native sheep breeds/populations, 35 breeds (see Supplementary Table 1; Dataset I) were genotyped with the Illumina Ovine SNP50 (54,241 SNPs) BeadChip, and 5 breeds (Hu, Wadi, Sishui Fur, Large-tailed Han and Small-tailed Han sheep; see Supplementary Table 1; Dataset II) were genotyped with the Illumina Ovine Infinium HD (685,734 SNPs) BeadChip. The Ovine SNP50 and Infinium HD BeadChips were developed by the International Sheep Genomics Consortium (ISGC; <http://www.sheephapmap.org>). Details regarding SNP discovery, the design of the ovine array and genotyping procedures for the BeadChips can be found at the following address: <http://www.sheephapmap.org/hapmap.php>, and in Kijas *et al.*³³ for the SNP50 BeadChip and Anderson *et al.*⁵⁰ for the HD SNP BeadChip.

Published Ovine SNP dataset. An Ovine SNP50 BeadChip dataset of 74 worldwide-distributed breeds/populations (Dataset III) was retrieved from a previous study³³. A set of 57 native breeds (Supplementary Table 1) that had been sampled and genotyped within the Sheep HapMap project were selected out of the 74 breeds (for information on the breeds and their geographic origins, see Kijas *et al.*³³).

Ovine SNP datasets quality control. All of the ovine SNP BeadChip datasets (Datasets I, II and III) were merged to apply quality control procedures using Plink v.1.9 software⁵¹. The SNP quality control measures have been detailed elsewhere, in Kijas *et al.*⁵², Miller *et al.*⁵³, Kijas *et al.*³³ and Lv *et al.*⁵⁴. Briefly, quality control was performed according to the following criteria: (1) individuals with a call rate >95%, (2) SNPs with a minor allele frequency (MAF) >0.05, (3) SNPs with a >95% genotyping rate, and (4) SNPs with physical locations on autosomes. After removing the SNPs and individuals that failed to meet any of the above criteria, 39,641 SNPs and 1,910 individuals were retained in our working dataset.

Bovine SNP datasets. Bovine SNP50 BeadChip datasets were retrieved from previous publications^{10,34}, which included 53 native cattle breeds across its main distributions in the world (Supplementary Table 2). SNPs and individuals that failed to meet any of the following criteria were removed: (1) individuals with a call rate >95%, (2) MAF > 0.05, (3) SNPs with a >95% genotyping rate, and (4) SNPs with physical locations on autosomes (see also Decker *et al.*³⁴). After filtering, there were 37,827 SNPs and 1,031 animals available for further analyses.

Genetic diversity. The observed (H_O) and expected (H_E) heterozygosity of each sheep and cattle breed were estimated using Plink v.1.9⁵¹. They are important within-breed genetic diversity parameters for the conservation priorities in domesticated animals. In addition, F_{IS} was calculated by genepop R package⁵⁵ based on the observed versus expected number of homozygous genotypes.

Effective population size. Effective population size (N_e) is a well-known parameter in evaluating conservation priorities for a species. Estimates of contemporary N_e were performed with the linkage disequilibrium method using NE ESTIMATOR v.2.1 software⁵⁶ with the following parameter settings: random mating model and a Pcrit value with 0.05 which as the criterion for excluding rare alleles. In order to insure all loci in the analysis are physically unlinked⁵⁷, we removed pairwise comparisons with $r^2 \geq 0.05$ before calculating N_e using Plink v.1.9 software⁵¹ (indep-pairwise 100 25 0.05). In addition, we divided the N_e value into three categories: $N_e \leq 100$, $100 < N_e < 200$, and $N_e \geq 200$.

Runs of homozygosity (ROH). Runs of homozygosity (ROHs) were detected using the Plink v.1.9 software⁵¹ for each individual. The following settings were generally suitable for ROH identification in the domesticated animals⁵⁸: (1) required minimum density (-homozyg-density 1000) and (2) number of heterozygotes allowed in a window (-homozyg-window-het 1). In this study, we detected ROH with the minimum length of 500 kb (-homozyg-window-kb 500) and a minimum SNP number of 50 SNPs (-homozyg-window-snp 50). The number (N_{ROH}) and average length (L_{ROH}) of ROH for each breed were estimated and the N_{ROH} and L_{ROH} were divided into three categories: $N_{ROH} \leq 100$, $100 < N_{ROH} < 200$, and $N_{ROH} \geq 200$; $L_{ROH} \geq 20$ Mb, $15 \text{ Mb} < L_{ROH} < 20 \text{ Mb}$, and $L_{ROH} \leq 15 \text{ Mb}$. In addition, the inbreeding coefficient based on ROH (F_{ROH}) was calculated for each animal according to the following formula^{26,59}:

$$F_{ROH} = \frac{\sum_k (Length(ROH_k))}{L} \quad (1)$$

where the numerator represents the sum of ROH per animal above 500 kb length and L is the total length of genome. Total lengths of the cattle and sheep genomes were considered as 2,510,611 kb⁶⁰ and 2,610,000 kb⁶¹, respectively. For sheep and cattle, the mean of F_{ROH} of each breed were estimated separately. The breed F_{ROH} values were divided into three ranges: $F_{ROH} \geq 0.10$, $0.05 < F_{ROH} < 0.10$, and $F_{ROH} \leq 0.05$.

Genetic structure analysis. In order to understand the relationship within and between breeds across each major geographic group, neighbor-joining (NJ) trees were constructed by the software SplitTree 4⁶². To avoid the SNPs with high levels of LD distorting the NJ tree, SNP were pruned using Plink v.1.9⁵¹. For the sheep and cattle dataset, the 39,641 and 37,827 SNPs were pruned to 34,770, and 29,093 SNPs respectively, by application of LD pruning command line-indep-pairwise 100 25 0.25. Subsequently, Reynolds genetic distances⁶³ were calculated by Arlequin 3.5.2.2⁶⁴, and the NJ trees were constructed using SplitsTree 4⁶² and visualized by FigTree v.1.3.1⁶⁵.

Analysis of conservation priorities based on the molecular coancestry approach. The method of molecular coancestry⁸ has proven to be efficient for defining conservation priorities for domestic animals. Unlike prioritization based on Weitzman diversity measures^{66,67}, this method considers within-breed genetic variation, which is of great importance for the management of livestock⁶⁸. The method assumes that there is a metapopulation including n populations with N_i breeding individuals. In the metapopulation, s_i denotes the average self-coancestry of the N_i individuals of population i ; D_{ij} is the average distance between individuals of the i^{th} population; and D_{ij} is Nei's minimum distance between populations i and j ⁶⁹. The average global coancestry (\bar{f}) was calculated as follows⁷⁰:

$$\bar{f} = \sum_{i=1}^n \frac{1}{n} \left[s_i - D_{ii} - \frac{\sum_{j=1}^n D_{ij}}{n} \right] \quad (2)$$

This equation can also be expressed based on genetic diversity⁷⁰:

$$(1 - \bar{f}) = (1 - \tilde{f}) + \bar{D} = 1 - \tilde{s} + \tilde{D} + \bar{D} = (1 - \tilde{s}) + (\tilde{s} - \tilde{f}) + (\tilde{f} - \bar{f}) \quad (3)$$

In equation (3), \tilde{f} and \tilde{s} are the mean coancestry and self-coancestry of the populations, respectively; \tilde{D} is Nei's minimum distance between individuals within populations; and \bar{D} is the average genetic distance over the entire metapopulation. In the derived expression, the total genetic diversity ($GD_T = 1 - \bar{f}$) was divided into three partitions: the genetic diversity within individuals ($GD_{WI} = 1 - \tilde{s}$), the genetic diversity between individuals ($GD_{BI} = \tilde{s} - \tilde{f}$) and the genetic diversity between populations ($GD_{BS} = \tilde{f} - \bar{f}$). The sum of the first two components (GD_{WI} and GD_{BI}) gives the genetic diversity within populations ($GD_{WS} = 1 - \tilde{f}$). As a result, the total genetic diversity GD_T is assumed to be the sum of the genetic diversity within populations (GD_{WS}) and the genetic diversity between populations (GD_{BS}), i.e., $GD_T = GD_{WS} + GD_{BS}$.

The loss or gain of genetic diversity (ΔGD_{WS} , ΔGD_{BS} and ΔGD_T) was estimated using Metapop v.2.0 software⁷¹ when each of the populations was removed from the dataset. The parameter of $\lambda = 1$, which indicates equal weight for within- and between-breed diversity, was applied in the Metapop calculations⁷¹ (Supplementary Tables 1 and 2). It is noteworthy that the initial version of Metapop used loss (-) vs gain (+) of diversity as indicative of high vs low contribution of diversity, respectively. However, in the latest version of Metapop (used in this paper), the sign was changed to go to a more intuitive view: + means greater contribution and - means lower contribution.

Analysis of conservation priorities based on the allelic richness approach. Allelic richness is an important measure of genetic diversity for setting conservation priorities in livestock. The rarefaction method developed by Hurlbert is advantageous for estimating allelic richness^{72,73}. With this method, El Mousadik and Petit denoted allelic richness as a_i , representing the number of different alleles in a sample of genes taken at random. The calculation formula is as follows⁷⁴:

$$a_i = \sum_{k=1}^K (1 - P_{ik}) = \sum_{k=1}^K \left(1 - \frac{\binom{N_i - N_{ik}}{g}}{\binom{N_i}{g}} \right) \quad (4)$$

where P_{ik} is the probability that allele k does not exist in a random sample of g genes; N_i represents the total number of genes in the sample of a given subpopulation i and N_{ik} represents the number of copies of the k th allele from that subpopulation.

The within-subpopulation component of allelic diversity was calculated according to the following formula⁷⁵:

$$A_s = \left(\frac{1}{n} \sum_{i=1}^n a_i \right) - 1 \quad (5)$$

The average distance between all subpopulations is

$$D_A = \frac{1}{n^2} \left[\sum_{i,j=1}^n d_{A,ij} \right] \quad (6)$$

where $d_{A,ij}$ is the average allelic distance between subpopulations i and j , and it is equal to

$$d_{A,ij} = \frac{1}{2} \sum_{k=1}^K [(1 - P_{ik})P_{jk} + P_{ik}(1 - P_{jk})] \quad (7)$$

The total allelic diversity is estimated with the following equation:

$$A_T = A_s + D_A = \left[\frac{1}{n} \sum_{i=1}^n \left(a_i + \frac{1}{n} \sum_{i=1}^n d_{ij} \right) \right] - 1 = \left[\frac{1}{n^2} \sum_{k=1}^K \sum_{i,j=1}^n (1 - P_{ik}P_{jk}) \right] - 1 \quad (8)$$

The three estimators (A_s , D_A and A_T) were calculated using Metapop v.2.0 software as described above (Supplementary Tables 1 and 2).

The results obtained by the molecular coancestry and allelic richness methods have some differences. However, we compared the estimated values based on both methods using paired t -tests, and found that they did not show significant differences for the three estimators (for sheep and cattle, ΔGD_{WS} : $t = 0.01$, $df = 96$, $P = 0.999$ and $t = 0.031$, $df = 52$, $P = 0.976$; ΔGD_{BS} : $t = 0.2$, $df = 96$, $P = 0.842$ and $t = 0.286$ $df = 52$, $P = 0.776$; and ΔGD_T : $t = 0.301$ $df = 96$, $P = 0.764$ and $t = 0.429$, $df = 52$, $P = 0.670$, respectively). In order to obtain more accurate results, the average values of the estimates obtained using the two approaches were employed in the following statistical analysis. The average values of ΔGD_{WS} , ΔGD_{BS} and ΔGD_T in sheep and cattle were divided into four fractions: < -0.1 , $-0.1-0$, $0-0.1$ and > 0.1 , and then plotted around the world using the software ArcGIS v.10.1 (ESRI Inc, Redlands, CA, USA) (Supplementary Fig. 4).

Analysis of the effect of regional economics on the conservation priority of different breeds.

We further investigated the associations between the changes (loss or gain) in genetic diversity for each breed (or group of breeds) and the economies of the countries (or regions) from which the breeds originated. The average GDP per capita of the countries (or regions) during the years 1993 to 2013 was collected from the World Bank database (<http://data.worldbank.org.cn/indicator>). Because the 40 native Chinese sheep populations originated from 13 provinces, which differ greatly in their economies and employ different breeding and conservation strategies for domestic animals⁷⁶, the GDP per capita for the provinces, rather than the whole country of China, was collected for the period from 1993 to 2013 (from the China Statistics Bureau; available at <http://data.stats.gov.cn>; Supplementary Table 5) and included in the analysis. The countries (or regions) were classified into three different categories based on their average GDP per capita (A-GDPPC) during these years: underdeveloped countries/regions ($A-GDPPC \leq \$1,500$); developing countries/regions ($\$1,500 \leq A-GDPPC \leq \$15,000$); and developed countries/regions ($A-GDPPC \geq \$15,000$). For sheep, the breeds came from 38 countries or regions: 9 were underdeveloped; 16 were developing; and 13 were developed. For cattle, the breeds came from 21 countries or regions: 10 were underdeveloped; 3 were developing; and 8 were developed (see Supplementary Tables 3 and 4). Subsequently, we calculated the loss or gain of diversity for each country/region when all of the breeds from the country/region were removed using the two approaches described above (Supplementary Tables 3 and 4). Differences in the loss or gain of genetic diversity averaged over sheep and cattle breeds in the three different economic categories were compared using the boxplots statistics in the boxplot package of *R* software v.3.4.3⁷⁷, and the P value was calculated using the Kolmogorov-Smirnov test⁷⁸.

To evaluate the effect of economic development on the loss or gain of genetic diversity for the breeds, we performed regression analysis between the loss or gain of genetic diversity (within subpopulation, ΔGD_{WS} ; between subpopulations, ΔGD_{BS} ; and total genetic diversity, ΔGD_T) and the increased average GDP per capita ($\Delta GDPPC$) for 1993 to 2013 from each country (or region) using the dual-regression approach. The coefficient of determination (R^2) and its significance (p -value) were calculated using SPSS v.22.0 software⁷⁹.

Data availability. The data sets analyzed in the present study are available from the corresponding author on reasonable request.

References

- Hodges, J. Jubilee history of the European Association for animal production: 1949–1999. *Livest. Prod. Sci.* **60**, 105–168 (1999).
- FAO. The second report on the state of the world's animal genetic resources for food and agriculture. In *FAO Commission on Genetic Resources for Food and Agriculture Assessments* (eds. Scherf, B. D. & Pilling, D) (Food & Agriculture Organization, Rome, 2015).
- Hoffmann, I., Boerma, D. & Scherf, B. The Global Plan of Action for Animal Genetic Resources—The road to common understanding and agreement. *Livest. Sci.* **136**, 7–14 (2011).

4. Rook, A. *et al.* Matching type of livestock to desired biodiversity outcomes in pastures—a review. *Biol. Conserv.* **119**, 137–150 (2004).
5. Lupton, C. J. ASAS Centennial Paper: Impacts of animal science research on United States sheep production and predictions for the future. *J. Anim. Sci.* **86**, 3252–3274 (2008).
6. Tisdell, C. Socioeconomic causes of loss of animal genetic diversity: analysis and assessment. *Ecol. Econ.* **45**, 365–376 (2003).
7. Drucker, A. G. & Anderson, S. Economic analysis of animal genetic resources and the use of rural appraisal methods: lessons from Southeast Mexico. *Int. J. Agr. Sustain.* **2**, 77–97 (2004).
8. Caballero, A. & Toro, M. A. Analysis of genetic diversity for the management of conserved subdivided populations. *Conserv. Genet.* **3**, 289–299 (2002).
9. Petit, R. J., El Mousadik, A. & Pons, O. Identifying populations for conservation on the basis of genetic markers. *Conserv. Biol.* **12**, 844–855 (1998).
10. Groeneveld, L. *et al.* Genetic diversity in farm animals—a review. *Anim. Genet.* **41**, 6–31 (2010).
11. The Bovine HapMap Consortium. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* **324**, 528–532 (2009).
12. Chessa, B. *et al.* Revealing the history of sheep domestication using retrovirus integrations. *Science* **324**, 532–536 (2009).
13. Lv, F.-H. *et al.* Mitogenomic meta-analysis identifies two phases of migration in the history of Eastern Eurasian sheep. *Mol. Biol. Evol.* **32**, 2515–2533 (2015).
14. Loftus, R. T., MacHugh, D. E., Bradley, D. G., Sharp, P. M. & Cunningham, P. Evidence for two independent domestications of cattle. *Proc. Natl. Acad. Sci. USA* **91**, 2757–2761 (1994).
15. McTavish, E. J., Decker, J. E., Schnabel, R. D., Taylor, J. F. & Hillis, D. M. New World cattle show ancestry from multiple independent domestication events. *Proc. Natl. Acad. Sci. USA* **110**, E1398–E1406 (2013).
16. Franklyn, I. R. Evolutionary change in small populations. In *Conservation Biology: An Evolutionary—Ecological Perspective* (eds Soulé, M. E. & Wilcox, B. A.), 135–149 (Sinauer Associates, Sunderland, MA, 1980).
17. Frankham, R., Bradshaw, C. J. & Brook, B. W. Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol. Conserv.* **170**, 56–63 (2014).
18. RBST. <http://www.rbst.org.uk/Our-Work/watchlist-overview> (2018).
19. Wiltshire Horn. https://en.wikipedia.org/wiki/Wiltshire_Horn (2018).
20. Coulson, T. *et al.* Age, sex, density, winter weather, and population crashes in Soay sheep. *Science* **292**, 1528–1531 (2001).
21. RBST. <http://www.rbst.org.uk/soay> (2018).
22. Talib, C., Entwistle, K., Siregar, A., Budiarti-Turner, S. & Lindsay, D. Survey of population and production dynamics of Bali cattle and existing breeding programs in Indonesia. In 'Strategies to improve Bali cattle in eastern Indonesia. ACIAR Proceedings 110' (eds Entwistle, K. & Lindsay, D.R.), 3–9 (Australian Centre for International Agricultural Research: Canberra, 2003).
23. Gibson, J., Morton, N. E. & Collins, A. Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.* **15**, 789–795 (2006).
24. Keller, M. C., Visscher, P. M. & Goddard, M. E. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics* **189**, 237–249 (2011).
25. Bosse, M. *et al.* Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. *PLoS Genet.* **8**, e1003100 (2012).
26. McQuillan, R. *et al.* Runs of homozygosity in European populations. *Am. J. Hum. Genet.* **83**, 359–372 (2008).
27. Ouborg, N. J., Pertoldi, C., Loeschcke, V., Bijlsma, R. K. & Hedrick, P. W. Conservation genetics in transition to conservation genomics. *Trends Genet.* **26**, 177–187 (2010).
28. Leroy, G. Inbreeding depression in livestock species: review and meta-analysis. *Anim. Genet.* **45**, 618–628 (2014).
29. Garcia-Gamez, E., Sahana, G., Gutierrez-Gil, B. & Arranz, J. J. Linkage disequilibrium and inbreeding estimation in Spanish Churra sheep. *BMC Genet.* **13**, 43 (2012).
30. Kim, E. S. *et al.* Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. *Plos One* **8**, e80813 (2013).
31. Kim, E. S., Sonstegard, T. S., Van Tassell, C. P., Wiggans, G. & Rothschild, M. F. The relationship between runs of homozygosity and inbreeding in Jersey cattle under selection. *Plos One* **10**, e0129967 (2015).
32. Ferencakovic, M. *et al.* Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *J. Anim. Breed. Genet.* **130**, 286–293 (2013).
33. Kijas, J. W. *et al.* Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *Plos Biol.* **10**, e1001258 (2012).
34. Decker, J. E. *et al.* Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *Plos Genet.* **10**, e1004254 (2014).
35. Hiemstra, S. J., Haas, Y., Mäki-Tanila, A. & Gandini, G. Local cattle breeds in Europe: Development of policies and strategies for self-sustaining breeds (Wageningen Academic Publishers, 2010).
36. Baylis, K., Peplow, S., Rausser, G. & Simon, L. Agri-environmental policies in the EU and United States: A comparison. *Ecol. Econ.* **65**, 753–764 (2008).
37. Gaspar, P., Mesías, F., Escribano, M. & Pulido, F. Assessing the technical efficiency of extensive livestock farming systems in Extremadura, Spain. *Livest. Sci.* **121**, 7–14 (2009).
38. Taberlet, P. *et al.* Are cattle, sheep, and goats endangered species? *Mol. Ecol.* **17**, 275–284 (2008).
39. Gandini, G. *et al.* Motives and values in farming local cattle breeds in Europe: a survey on 15 breeds. *Anim. Genet. Resour.* **47**, 45–58 (2010).
40. Blackburn, H. Genebank development for the conservation of livestock genetic resources in the United States of America. *Livest. Sci.* **120**, 196–203 (2009).
41. Samson, F. & Knopf, F. Prairie conservation in north america. *BioScience* **44**, 418–421 (1994).
42. Åby, B., Aass, L., Sehested, E. & Vangen, O. Effects of changes in external production conditions on economic values of traits in Continental and British beef cattle breeds. *Livest. Sci.* **150**, 80–93 (2012).
43. Hiemstra, S. *et al.* What's on the menu? Options for strengthening the policy and regulatory framework for the exchange, use and conservation of animal genetic resources. *Anim. Genet. Resour.* **41**, 65–74 (2007).
44. Gollin, D., Van Dusen, E. & Blackburn, H. Animal genetic resource trade flows: Economic assessment. *Livest. Sci.* **120**, 248–255 (2009).
45. Hoffmann, I. Livestock biodiversity. *Rev. Sci. Tech. Off. Int. Epiz.* **29**, 73–86 (2010).
46. Hoffmann, I. Climate change and the characterization, breeding and conservation of animal genetic resources. *Anim. Genet.* **41**, 32–46 (2010).
47. Huo, J. *et al.* Genetic diversity of local Yunnan chicken breeds and their relationships with Red Junglefowl. *Genet. Mol. Res.* **13**, 3371–3383 (2014).
48. López Herráez, D. *et al.* Genetic variation and recent positive selection in worldwide human populations: evidence from nearly 1 million SNPs. *Plos One* **4**, e7888 (2009).
49. Sambrook, J. & Russel, D. W. *Molecular Cloning: A Laboratory Manual 3rd edn* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2001).
50. Anderson, R. *et al.* Development of a High Density (600 K) Illumina Ovine SNP Chip and Its Use to Fine Map the Yellow Fat Locus. In *Plant and Animal Genome XXII Conference* (Plant and Animal Genome, San Diego, 2014).

51. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
52. Kijas, J. W. *et al.* A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. *PLoS One* **4**, e4668 (2009).
53. Miller, J. M., Poissant, J., Kijas, J. W. & Coltman, D. W. & International Sheep Genomics Consortium. A genome-wide set of SNPs detects population substructure and long range linkage disequilibrium in wild sheep. *Mol. Ecol. Res.* **11**, 314–322 (2011).
54. Lv, F.-H. *et al.* Adaptations to climate-mediated selective pressures in sheep. *Mol. Biol. Evol.* **31**, 3324–3343 (2014).
55. Rousset, F. Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103–106 (2008).
56. Do, C. *et al.* NEESTIMATOR v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol. Ecol. Resour.* **14**, 209–214 (2014).
57. Waples, R. K., Larson, W. A. & Waples, R. S. Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity* **117**, 233–240 (2016).
58. Zhao, Y.-X. *et al.* 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**, 2380–2395 (2017).
59. Karimi, K., Koshkoiyeh, A. E., Fozzi, M. A., Porto-Neto, L. R. & Gondro, C. Prioritization for conservation of Iranian native cattle breeds based on genome-wide SNP data. *Conserv. Genet.* **17**, 77–89 (2016).
60. Purfield, D. C., Berry, D. P., McParland, S. & Bradley, D. G. Runs of homozygosity and population history in cattle. *BMC Genet.* **13**, 70 (2012).
61. Jiang, Y. *et al.* The sheep genome illuminates biology of the rumen and lipid metabolism. *Science* **344**, 1168–1173 (2014).
62. Perl, Y. Optimum split trees. *J. Algorithms* **5**, 367–374 (1984).
63. Reynolds, J., Weir, B. S. & Cockerham, C. C. Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**, 767–779 (1983).
64. Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
65. Rambaut, A. & Drummond, A. FigTree v1.3. 1. *Institute of Evolutionary Biology, University of Edinburgh* (2010).
66. Weitzman, M. L. On diversity. *Q. J. Econ.* **107**, 363–405 (1992).
67. d'Arnoldi, C. T., Foulley, J.-L. & Ollivier, L. An overview of the Weitzman approach to diversity. *Genet. Sel. Evol.* **30**, 149–161 (1998).
68. Ollivier, L. & Foulley, J.-L. Aggregate diversity: new approach combining within-and between-breed genetic diversity. *Livest. Prod. Sci.* **95**, 247–254 (2005).
69. Cockerham, C. C. Group inbreeding and coancestry. *Genetics* **56**, 89–104 (1967).
70. Fabuel, E., Barragán, C., Silió, L., Rodríguez, M. & Toro, M. Analysis of genetic diversity and conservation priorities in Iberian pigs based on microsatellite markers. *Heredity* **93**, 104–113 (2004).
71. Pérez-Figueroa, A., Saura, M., Fernández, J., Toro, M. A. & Caballero, A. METAPOPOP—A software for the management and analysis of subdivided populations in conservation programs. *Conserv. Genet.* **10**, 1097–1099 (2009).
72. Sanders, H. L. Marine benthic diversity: a comparison study. *Am. Nat.* **102**, 243–282 (1968).
73. Hurlbert, S. H. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* **52**, 577–586 (1971).
74. El Mousadik, A. & Petit, R. J. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theor. Appl. Genet.* **92**, 832–839 (1996).
75. Caballero, A. & Rodríguez-Ramilo, S. T. A new method for the partition of allelic diversity within and between subpopulations. *Conserv. Genet.* **11**, 2219–2229 (2010).
76. Huang, D., Wang, K. & Wu, W. Problems and strategies for sustainable development of farming and animal husbandry in the Agro-Pastoral Transition Zone in Northern China (APTZNC). *Int. J. Sust. Dev. World Ecol.* **14**, 391–399 (2007).
77. R Development Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna* (2018).
78. Gail, M. H. & Green, S. B. A generalization of the one-sided two-sample Kolmogorov-Smirnov statistic for evaluating diagnostic tests. *Biometrics* **32**, 561–570 (1976).
79. Mather, L. E. & Austin, K. L. The Statistical Package for the Social Sciences (SPSS) as an adjunct to pharmacokinetic analysis. *Biopharm. Drug Dispos.* **4**, 157–172 (1983).

Acknowledgements

This work was supported by the grants from the National Natural Science Foundation of China (grants Nos 91731309 and 31661143014).

Author Contributions

J.Y. and F.-H.L. conceived and designed the project; M.Z. and W.-F.P. analysed the data with contributions from X.-J.H., F.-H.L. and Y.-X.Z.; M.Z. and W.-F.P. wrote the paper with contribution from J.Y. All authors reviewed and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-30061-0>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018