

# Genomic Diversity in Pig (*Sus scrofa*) and its Comparison with Human and other Livestock

Chunyan Zhang and Graham Plastow\*

1400 College Plaza, Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB T6G 2C8, Canada

**Abstract:** We have reviewed the current pig (*Sus scrofa*) genomic diversity within and between sites and compared them with human and other livestock. The current Porcine 60K single nucleotide polymorphism (SNP) panel has an average SNP distance in a range of 30 - 40 kb. Most of genetic variation was distributed within populations, and only a small proportion of them existed between populations. The average heterozygosity was lower in pig than in human and other livestock. Genetic inbreeding coefficient ( $F_{IS}$ ), population differentiation ( $F_{ST}$ ), and *Nei's* genetic distance between populations were much larger in pig than in human and other livestock. Higher average genetic distance existed between European and Asian populations than between European or between Asian populations. Asian breeds harboured much larger variability and higher average heterozygosity than European breeds. The samples of wild boar that have been analyzed displayed more extensive genetic variation than domestic breeds. The average linkage disequilibrium (LD) in improved pig breeds extended to 1 - 3 cM, much larger than that in human (~ 30 kb) and cattle (~ 100 kb), but smaller than that in sheep (~ 10 cM). European breeds showed greater LD that decayed more slowly than Asian breeds. We briefly discuss some processes for maintaining genomic diversity in pig, including migration, introgression, selection, and drift. We conclude that, due to the long time of domestication, the pig possesses lower heterozygosity, higher  $F_{IS}$ , and larger LD compared with human and cattle. This implies that a smaller effective population size and less informative markers are needed in pig for genome wide association studies.

Received on: December 30, 2010 - Revised on: January 21, 2011 - Accepted on: January 22, 2011

**Keywords:** Pig, genomic diversity, linkage disequilibrium, human, livestock.

## 1. INTRODUCTION

Worldwide, domestic pigs can be divided into two main clades: Asian- and European-types, which diverged from each other around 58,000 years ago. The European wild boars were domesticated at least in the 4th millennium BC and then rapidly spread throughout Europe [1, 2]. Asian pigs were thought to have been introduced into Europe during the late 18th and early 19th centuries [3], and recent studies suggest this has had a significant impact on the diversity of these breeds. Significant breed development took place from the 18th Century, followed by extensive organized genetic improvement (artificial selection) through the application of quantitative genetics theory in the second half of the 20<sup>th</sup> Century. Even so a large number of local breeds are maintained especially in China. It is commonly held that current European and Asian pig breeds are domesticated from different ancestors that might have different genome wide diversity and extent of linkage disequilibrium (LD).

Initial genetic studies in the pig mainly focused on investigating the genetic diversity of a breed at relatively small numbers of individual sites, which is essential for sustainable management of genetic resources for future utility. With the

identification of the first useful DNA marker (HAL1843) in the ryanodine receptor gene for marker assisted selection [4], a burst of research looking for variation in DNA sequence took place in order to explain useful variation in economic traits. The availability of many genetic markers enabled the search for quantitative trait loci (QTL) using various breeding populations [5, 6] or by candidate gene approaches [7]. Typically, QTL studies made use of the higher variability of microsatellite markers and the generation of their maps. Candidate gene studies resulted in the identification of SNPs or indels. One of the key aspects for utility in marker assisted selection was the identification of consistent effects across different breeding populations or lines [8-10]. This increased confidence in the associations and simplified application. This across population marker relationship was essentially searching for markers that are in LD with causative mutations [11]. This desire for LD across populations meant that true associations within an individual line were often not used (type-2 error), but this was made up for by their more widespread utility. As the identification of large numbers of SNPs became easier then it was possible to explore the extent of LD in livestock, in order to determine the density required for different applications especially genome wide association studies. The availability of dense SNP panels (>50,000 SNPs per species) enables many different marker studies. Applications in the pig, include: inference on population history, structure and dynamics, estimation of effective population size, QTL mapping strategies, and whole genome association studies and genomic selection [12,13].

\*Address correspondence to this author at the 1400 College Plaza, Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB T6G 2C8, Canada; Tel: 780- 492-1496; Fax: 780-248-1900; E-mail: graham.plastow@ales.ualberta.ca

Therefore, knowledge of genomic diversity within and between populations is becoming more and more important in current livestock research.

The development of high throughput genomic techniques provides us with an opportunity to assess genome wide diversity and its structure. Since the earlier draft sequence published for human (2001) and cattle (2009), an increasing number of studies have aimed at studying genomic diversity and LD in human and cattle [12-16]. For example, panels of more than 50,000 SNPs are available for cattle, pigs, sheep, chickens and horses and the latest panels for cattle contain more than 800,000 SNPs. The pig is now catching up these other species. To date most studies are based on a low marker density or a limited genomic region, which limits the utility of the results [1, 17-22]. The development of the Porcine 60K SNP panel helps to investigate genome wide diversity with a higher resolution [23]. Compared with other livestock, porcine populations are heavily structured and arranged in breeds for economic interest and environments. Comparative genomic diversity enables us to explore the degree of genomic variation and LD among pig and other species. This also helps to detect genomic regions that have been subject to selective sweeps in different pig populations.

The focus of this review is on pig, however, we will compare the pig with results from human as well as other livestock species. We begin by discussing pig genomic diversity within sites and then addressed the genomic diversity among sites described by LD. In each case, we briefly compare pig genomic diversity with human and other livestock. The mechanisms for maintaining genomic diversity in pig are briefly discussed, including the factors of migration, introgression, selection and drift. A comprehensive review on population genomic diversity from the perspective of evolutionary theories is detailed in this volume by Hu *et al.* [24].

## 2. PIG GENOMIC DIVERSITY WITHIN SITES

### 2.1. Genetic Diversity within Populations

Pig genomic diversity within populations is quite variable. Observed heterozygosity ranged from 0.35 to 0.60, with an average  $\sim 0.5$  across 17 autosomal chromosomes in 11 European pig breeds [25]. These values were similar to those in Chinese populations, ranging from 0.429 to 0.677 [19-20]. The expected heterozygosity was much higher than the observed heterozygosity in many reports. The average expected heterozygosity ranged from 0.53 to 0.80 among 13 populations from both Asia and Europe, including domestic and wild boars [20]. Chinese population had much higher diversity, ranging from 0.700 to 0.876 from 18 Chinese pig breeds [26]. If we consider a worldwide scale, pig populations had quite variable heterozygosities on different chromosomal regions. For example, the mean heterozygosity varied from 0.56 to 0.68 on chromosome 4 (SSC4) and from 0.65 to 0.80 on chromosome 7 (SSC7) in different populations [21]. This may reflect the relatively long time of breeding and selection for the pig.

A large number of SNPs are currently available for most species, such as human [27], cattle [15] and pig [23]. Recently, more than 372,000 SNPs were identified in swine, among which 45,510 SNPs were mapped to specific chromosomes in porcine genome build 7, including 21 SNPs on

chromosome Y [23]. This number has now grown to more than 23 million (M. Groenen personal communication) with around 9M shared by both European and Chinese breeds. The number of SNPs varies largely on different chromosomes. With the exception of chromosome Y, most SNPs were identified on chromosome 1, followed by chromosomes 14, 4, and 7. Chromosome 12 had the least number of SNPs. This distribution partly reflects the cytogenetic map of the pig – chromosome 1 is the largest metacentric chromosome and 12 the smallest. The ratio of transitions to transversions was 2:1 [28-30], similar to that observed in human [31].

These SNPs were used to construct a panel using the Illumina platform, 64,232 SNPs were decoded and selected [23]. This panel shows quite a variable marker distribution among chromosomes and between genome builds as well as a proportion of unmapped SNPs, reflecting the incomplete state of knowledge of the pig genome to date. The average distance between SNPs on porcine genome build 7 was 30-40 kb, except for the X chromosome that had the larger distance (59.2 kb). This is close to the target required for efficient genome wide association studies according to estimates of LD. However, the largest distance between SNPs was around 450 kb on chromosomes 13 and 15, and the smallest gap was 161.2 kb on chromosome 18. The largest distances between SNPs in build 8 were generally higher than those in build 7. The number of intervals with a gap size larger than 250 kb was 115 and 207 in builds 7 and 8, respectively. In genome build 7, most gaps were on chromosomes 14 and X. Some chromosomes (2, 3, 6, 10, 12, 16 and 18) did not have any large gaps between the SNPs. In build 8, an increasing number of large gaps were observed for all chromosomes except for SSC14 where unmapped SNPs in build 7 were added in build 8, filling the observed large gaps [23]. These differences reflect the maturation of the genome sequence with different builds correcting some and introducing other errors. The latest build of the swine genome is build 10 which will be used for preparation of the sequence publication [32].

### 2.2. Genetic Diversity Among Populations

Since pig breeds from Asian and European populations were domesticated from different ancestors, they are expected to have different genetic diversities. One suggestion is that Asian wild boars had higher genetic diversity compared to European wild boars [2]. The nucleotide diversity in terms of Watterson's method ( $\theta$ ) was 30% larger in Asian than in European populations [22]. Asian populations also have higher average heterozygosity than European populations. The observed and expected heterozygosities for worldwide breeds are summarized in Table 1 [19-20, 25-26]. It can be viewed that the average observed heterozygosity in Asian breeds was 0.566, ranging from 0.332 to 0.702, which in European breeds was around 0.542, at a range of 0.35 - 0.65. Furthermore, the average expected heterozygosity was also higher for Asian breeds (0.752) than for European breeds (0.570). Compared to domestic breeds, wild boars have a much higher observed heterozygosity, with an average of 0.628 and the range of 0.55 - 0.68.

The genetic distance among populations is very varied reflecting the differences in domestication model and history

**Table 1. Summary of Observed and Expected Heterozygosity Estimates in European and Asian Breeds as well as Wild Boars Adapted from the Literature [19, 20, 25, 26]**

	Observed Heterozygosity ( $H_o$ )			Expected Heterozygosity ( $H_e$ )		
	European	Asian	Wild	European	Asian	Wild
No of observations	17	28	5	17	28	5
Range	0.35 - 0.65	0.33 - 0.70	0.55 - 0.68	0.35 - 0.71	0.43 - 0.89	0.66 - 0.76
Average	0.542	0.566	0.628	0.570	0.752	0.708
S.D	0.078	0.087	0.050	0.095	0.132	0.039

in the two regions. We have summarized the *Nei's* genetic distances between populations in Table 2 [19, 25, 26]. The average genetic distance in Asian populations was slightly higher than that in European populations. In Asian breeds, the average *Nei's* genetic distance was 0.531, with a range from 0.194 to 1.188. In European populations, the average genetic distance was 0.495, ranging from 0.163 to 1.122. Higher average genetic distance was observed between European and Asian clades (1.434), which was nearly three times larger than those within populations. The smallest value was even up to 1.119 between Duroc and Tibet pigs [26]. However, Kim *et al.* [1] reported a smaller value (~ 0.017) between Asian and European pigs based on mtDNA D-loop sequences using the maximum-likelihood method, which was also much larger than that within Asian and European populations (~ 0.004).

Most genetic variation occurred within populations rather than between breeds from Asia and Europe. A summary of  $F_{ST}$  from the literature [19, 25, 26] is provided in Table 2. The average  $F_{ST}$  value (0.257) between Asian and European populations was much higher than those within populations. When compared within the two different regions, the overall genetic differentiation for Asian populations was ~ 0.227, ranging from 0.182 to 0.294, much higher than that for European (0.134). The  $F_{ST}$  value between domestic and wild boars for Asian (0.29) was also higher than that for European (0.194) [19, 25, 33]. A recent study using microsatellite markers also confirmed that the largest pairwise  $F_{ST}$  was between Asian and European breeds (0.410, between Meishan and Hampshire), and the smallest was observed within European breeds (0.021, between Duroc and Yorkshire) [20].

All Asian and European pigs were closely related in terms of their maternal lineages, but they were different from

each other. In wild boars, the average percentage sequence divergences calculated by mtDNA were 0.2370 within Chinese population, 0.3718 within European, and 0.5603 between Chinese and European. In domestic populations, the pairwise nucleotide sequence divergence was much lower (0.0056) between Chinese and European breeds [34]. Asian domestic populations had much larger variability than European breeds. Analysis based on mtDNA markers in 1536 samples (45 European and 21 Chinese breeds) indicated that the average frequency of mtDNA haplotypes in Asian was 29% of European breeds, but varied from 0 to 100% within individuals. A total of 28 Asian haplotypes were found in Chinese pigs, but only 6 Asian haplotypes were shared between European and Chinese populations [35]. These differences in genetic diversity are consistent with a strikingly different population history of humans and domestic animals. One separate study also revealed that the frequency of Asian haplotypes was low or absent in Duroc and Hampshire lines. The Landrace lines were less affected by Asian introgression than Large White lines. The haplotype in Pietrain which was originally developed in Belgium was completely absent in German and some commercial breeding lines, but showed a very high frequency of Asian haplotypes in some French lines [35]. It is worthwhile to note that Chinese breeds attributed high genetic diversity, with a higher level of haplotype diversity and smaller haplotype blocks than both wild boars and European breeds, but shared high levels of frequent haplotypes with Large White, Landrace, and Duroc. They also had a lower percentage of SNPs with the minor allele frequency (MAF) < 0.05. For instance, Meishan, a famous prolific Chinese breed, shared haplotypes that occurred at high frequency in European breeds [17].

**Table 2. A Summary of Genetic Distance (*Nei*) and Genetic Variation ( $F_{ST}$ ) between Different Populations from the Published Literature [19, 25, 26]**

	<i>Nei</i>			$F_{ST}$		
	Within European	Within Asian	Between Asian and European	Within European	Within Asian	Between Asian and European
No of observations	58	174	52	10	3	15
Range	0.163 - 1.122	0.194 - 1.188	1.119 - 1.794	0.021 - 0.209	0.182 - 0.294	0.169 - 0.410
Average	0.495	0.531	1.434	0.134	0.227	0.257
S.D	0.189	0.188	0.166	0.055	0.059	0.080

### 2.3. A Comparison with Human and Other Livestock

A comparison of pig genomic diversity with human and other livestock is summarized in Table 3. It shows that heterozygosity in pig was smaller than that in human and other livestock, but more flexible. Genetic variation within populations and genetic distance were much larger in pig than in human and other livestock.

The average expected heterozygosity in human ranged from 0.7 to 0.9 [36], which was higher than the values observed in pig (~0.5). Other primates had lower heterozygosity than those in pig, such as 0.38 observed in gorilla [37] and 0.32 observed between Western and Eastern chimpanzee [38]. The genetic diversity between populations ( $F_{ST}$ ) was lower in human than in pig (0.021 - 0.410, Table 3). In human,  $F_{ST}$  was 0.05 - 0.13 for autosomal SNPs [39-43], ~0.11 for autosomal copy number variations (CNVs) in a small set of populations, and 0.09 - 0.10 for *Alu* insertions polymorphism [43, 44]. Biswas *et al.* [14] indicated that 85% - 95% of human genetic variation was attributable to differences among individuals (higher heterozygosity), and that 5% - 15% was due to the differences between populations (lower  $F_{ST}$ ).

The genetic diversity was more extensive in goat than in pig (Table 3). Recent studies [45, 46] using more than five southern Indian goat breeds based on microsatellite markers showed that the genetic diversity within breeds ( $F_{IS}$ ) ranged from 0.03 to 0.32. The observed and expected heterozygosities were at ranges of 0.42 - 0.67 and 0.61 - 0.73, respectively. Among different goat breeds, *Nei's* genetic distance differed from 0.067 to 0.830. The genetic diversity between breeds ( $F_{ST}$ ) ranged from 0.012 to 0.200. The overall proportion of genetic differentiation among breeds was around 0.14, such as 0.02 in Guadarrama goat [47], 0.06 and 0.17 in Egyptian and Italian goat breeds [48], 0.13 in north-western goat breeds of India [46], and 0.15 in indigenous goats of Sub-Saharan Africa [49]. Compared to pig, the average proportion of genetic diversity within goat breeds is much larger, and among breeds is much lower.

The average heterozygosity in cattle is slightly higher than that in pig (see Table 3). The observed and expected heterozygosities in cattle were at the ranges of 0.47 - 0.74 and 0.45 - 0.78, respectively. While the genetic diversities both within ( $F_{IS}$ ) and between breeds ( $F_{ST}$ ) seem to be a slightly lower in cattle than in pig. *Nei's* genetic distance was much larger in pig (0.163 - 1.794) than in cattle (0.015 - 0.382). Compared to pig, cattle genetic diversity has been

more widely studied. Genetic diversity in Chinese breeds was higher than that in European breeds. This is similar to that found for the pig. In Chinese cattle breeds, the mean heterozygosity was at a range of 0.69 - 0.76, and *Nei's* genetic distance ranged from 0.025 to 0.352 [50]. In European breeds, the observed and expected heterozygosities were 0.49 - 0.72 and 0.45 - 0.71, respectively. The genetic distance was at a range of 0.029 - 0.309 [51, 52]. In Brazilian breeds, the average observed and expected heterozygosity were much higher, ranging from 0.6316 to 0.7409 and 0.7151 to 0.7839, respectively. Genetic diversity within ( $F_{IS}$ ) and between ( $F_{ST}$ ) populations were slightly lower. *Nei's* distance was at a range of 0.084 - 0.382 [53]. Pakistan cattle breeds displayed a moderate observed (0.47 - 0.51) and expected heterozygosity (0.63 - 0.67) [54].

## 3. PIG GENOMIC DIVERSITY AMONG SITES

### 3.1. LD within Populations

Within pig populations, a wide range of LD was observed on different chromosomes. With 29 and 5 microsatellite markers located on SSC15 and SSC2 in different lines, respectively, LD in terms of  $r^2$  was higher for SSC2 (0.35 - 0.48) than for SSC15 (0.15 - 0.19) [55]. By investigating LD on SSC4 and SSC7 in five domestic pig populations, the mean  $D'$  was higher on SSC7 than on SSC4, and LD decreased faster with the marker distance on SSC4 than on SSC7 [21]. Using the Porcine 60K SNP Beadchip, chromosomes 10 and 12 had the lowest average LD, chromosomes 1, 13, and 14 had the highest value [56]. Using the 60K panel, at the distance of 30 kb interval, the average LD on different autosomal chromosomes varied from 0.39 to 0.55 ( $r^2$ ). When considering longer SNP pair distance (3 Mb), differences in LD between different chromosomes were even larger up to 0.2 (for  $r^2$  value) [56]. It indicated that LD distance in pig could extend to 3 Mb even with a higher value of  $r^2$ . The significant LD variation among different regions was mainly due to the strong artificial selection on economic traits.

$D'$  and  $r^2$  are widely used to evaluate inter-site associations in livestock.  $D'$  strongly depends on allele frequency, and it decreases as the MAF increases.  $r^2$  is more influenced by linkage distance. By using ~4,500 autosomal SNPs in more than 6,000 pigs from six commercial lines [18], the highest average  $D'$  was observed in case of MAF as 0.067. The  $r^2$  values indicated that the largest LD was in the most tightly linked group. As the linkage distance increased from

**Table 3. Comparison of Pig Genomic Diversity with Human and other Livestock Summarized from Published Results**

Species	Within Populations			Between Populations		References
	$H_o$	$H_e$	$F_{IS}$	<i>Nei</i>	$F_{ST}$	
Pig	0.33 - 0.70	0.35 - 0.88	0.15 - 0.52	0.163 - 1.794	0.077 - 0.270	[19, 20, 25, 26]
Human	-	0.7 - 0.9	-	-	0.027 - 0.160	[14, 27, 75]
Cattle	0.47 - 0.74	0.45 - 0.78	0.05 - 0.26	0.015 - 0.382	0.028 - 0.216	[50-54]
Goat	0.42 - 0.67	0.61 - 0.73	0.03 - 0.32	0.067 - 0.830	0.012 - 0.200	[45, 46]

$H_o$  and  $H_e$ : observed and expected heterozygosity, respectively.

$F_{IS}$  and  $F_{ST}$ : fixation indices of genetic diversity within and between populations.

*Nei*: genetic distances between populations estimated by *Nei* standard method.

0.1 to 40 cM, the average  $r^2$  reduced from 0.371 to 0.008. The observed average  $r^2$  began to rapidly decrease when the linkage distance approached to 3 cM. Amaral *et al.* [17] used high density SNP markers across 18 chromosomes in more than 20 pig breeds to evaluate pig LD. They suggested that LD extended to 1 - 3 cM in these pig populations, with  $r^2$  being 0.3 as a threshold. Marker density and sampling size are the main factors that affect the accuracy of LD evaluation, which further influences the design of whole genome wide association studies. Nsengimana *et al.* [21] suggested that a marker density of 5 - 10 cM according to  $D'$  value was feasible in commercial pig populations for genome-wide association. Du *et al.* [18] recommended a considerably higher density of 0.1 to 1 cM for an initial whole genome scan. After a comparison among the distances from 40 - 60 kb [57] up to 400 kb [17] with commercial pig breeds, Ramos *et al.* [23] predicted that a density of 5 - 10 markers per cM was needed to conduct whole genome association studies in European breeds. However, a higher density is required for Asian breeds (see below).

**3.2. LD AMONG POPULATIONS**

The genome wide inter-site associations in terms of LD differ among different populations. Modern breeding programs increased the extent of LD and caused significant differences between European and Chinese pig breeds. LD decayed more rapidly in Chinese breeds than in European breeds, indicating that the extent of LD was smaller in Chinese breeds than in European breeds. In European breeds, LD extended up to 2 cM, and large haploblocks could be up to 400 kb; whereas in Chinese breeds, the extent of LD was smaller (0.05 cM) and generally did not exceed 10 kb [17, 58]. Chinese breeds had smaller LDs than European counterparts [17]. LD in Meishan breeds was larger than other Chinese breeds, but still smaller than that in most European breeds. When the whole genome was considered, the SNP spacing for European pig breeds was 0.1 cM, and 30,000 SNPs per individual were informative (with a MAF < 0.05 and  $r^2 = 0.3$ ). However, for Chinese breeds with a similar sample size, the SNP spacing was 0.005 cM, and 500,000 SNPs per individual would be required [17]. This large difference in LD between European and Chinese breeds can be explained by the different ancestral stocks and modern breeding systems that produced smaller effective population sizes in Europe. However, both Chinese and European populations were not significantly different from wild boars in LD. The European wild boars showed an intermediate LD between Chinese and European domestic breeds [17].

Contrary to Chinese breeds, European breeds showed significant differences in LD [17]. By using the Porcine 60K SNP BeadChip [56], at the SNP distance interval of 5 Mb in Finnish breeding populations, Landrace displayed lower LD than Yorkshire. In Landrace, the percentages of adjacent SNP pairs with  $r^2 > 0.3$  and  $r^2 > 0.2$  were 49% and 57%, respectively; while the corresponding percentages for Yorkshire were 52% and 60%. At the average useful  $r^2 > 0.2$ , the SNPs density in Landrace and Yorkshire extended to 1.0 and 1.5 Mb, respectively, which resulted in a smaller effective population size ( $N_e$ ) for Yorkshire (55) than for Landrace (80). Similar  $N_e$  were also obtained in American Landrace (74) and Berkshire (77), but higher in Hampshire (109) and Duroc (113) [59]. This, to some extent, reflected the different LDs in various populations. The variation of  $D'$  was also highly significant in Yorkshire, Large White and Landrace populations. The lowest  $D'$  was in Large White breeds, and the highest was in Duroc breeds [21].

**3.3. Comparison of Pig LD with Human and other Livestock**

A comparison of LD among pig, human, and other livestock is summarized in Table 4. It shows that, at the same  $r^2$  threshold, LD was significantly larger in pig than in human, but much smaller than in sheep. Compared with cattle, LD was much larger in pig at threshold of  $r^2 = 0.2$  but similar to cattle at threshold of  $r^2 = 0.1$ .

**3.3.1. Human**

Due to a smaller effective population size and the stronger selection that has occurred in livestock, LD was significantly greater in pig (1 - 3 cM) than in human. The average LD in human was up to ~ 30 kb with a high variability, depending on the different populations and marker density used. Computer simulations and empirical data suggested that human LD extended only 3 - 5 kb for common disease SNPs, meaning that approximately 500,000 SNPs at least was needed for whole-genome studies [60, 61]. Later studies revealed that LD could extend to a distance greater than 100 kb in some cases [62, 63]. Among different populations, it revealed that LD in northern European populations was at a range of 10 - 30 kb, while LD in northern African populations was much lower [64]. Reich *et al.* [65] analyzed various extents of LD in different chromosomes based on common alleles and found levels of LD extending up to 40 - 160 kb in different regions. The United States population of north-European descent displayed LD extending 60 kb, which was greater than that in a Nigerian population.

**Table 4. Comparison of Pig LD Ranges with Human and other Livestock Summarized from Published Results**

Species	$r^2$ range		References
	> 0.2	> 0.1	
Pig	< 3 Mb (cM)	< 5 Mb (cM)	[17, 18, 56]
Human	< 30 kb	< 100 kb	[64, 66]
Cattle	< 100 kb	< 5 cM	[12, 13, 15, 67, 68, 70]
Sheep	< 10 cM	< 30 cM	[73]

$r^2$ : LD estimated by  $r^2$  value.

The LD information is very useful for estimating  $N_e$  in different populations (see above for examples in pigs). For the SNP-pair interval less than 100 kb across all of the genome, the  $N_e$  estimate was ~3100 for western European, Japanese and Chinese, and ~7500 for the Yoruba population [66]. This result supports the out-of-Africa theory of ancestral human population expansion and concurrent bottlenecks. Compared with human, pig has a significantly larger extent of LD (1 - 3 cM) and results in a smaller  $N_e$  and fewer informative markers required for genome wide studies, at least for the European pig populations.

### 3.3.2. Cattle

Generally, cattle have larger LD than human, but smaller LD than pig. LD in cattle was at a range of 40 - 100 kb at the threshold of  $r^2 > 0.2$  [12, 15, 67]. de Roos *et al.* [12] compared 2430 animals from 5 breeds genotyped for 3072 SNP markers crossing 30 chromosomes. The average  $r^2$  was 0.14 at marker distance of 100 kb. It indicated that at least 50,000 SNP markers were required for whole genomic selection. McKay *et al.* [13] reported the similar extensive LD by approximately 2670 genome markers within eight cattle breeds. They displayed that the average  $r^2$  was 0.15 - 0.2 at a physical distance of 100 kb SNP loci apart, indicating that 30,000 - 50,000 markers were needed to conduct whole genome association studies. When microsatellite markers were used with four cattle breeds to estimate LD [16], at a marker-pair distance less than 5 cM, the average  $r^2$  value across the populations was 0.16 at a range of 0.11 - 0.22. For the distance greater than 50 cM, the average  $r^2$  declined to 0.07. Compared with cattle, pig LD was much stronger, especially for the useful  $r^2 > 0.2$ .

Extensive LD in cattle varied a little among different populations. In Holstein-Friesian cattle, by analyzing 1,566,890 syntenic and 365,400 non-syntenic SNP pairs that cover all autosomes [68], the significantly useful LD extended to 40 kb for  $r^2$  and 8.2 Mb for  $D'$ . It indicated that at least 75,000 SNPs and a sample size of 75 or 400 would be required for whole genome association study. Qanbari *et al.* [15] used Illumina Bovine 50K SNP BeadChip to analyse LD structure in German Holstein-Friesian cattle. A mean value of 0.21 ( $r^2$ ) was observed for SNPs less than 100 kb apart. For American Holstein cattle,  $r^2$  was much larger (0.59) at the same marker distance (100 kb) [69]. Kim & Kirkpatrick [70] also revealed strong LD ( $r^2 > 0.8$ ) in genomic regions with marker distance less than 50 kb. LD for SNP pair intervals 100 kb apart ( $r^2 = 0.14$ ) was similar to that in Holstein-Friesian cattle. Compared with pig, LD among populations varied less in cattle. This is in agreement with the common notion that a larger proportion of genetic variation exists within than between populations in most organisms.

### 3.3.3. Sheep

Compared with pig, LD in sheep seems much more extensive and flexible. High LD extending to 10 - 60 cM was observed in two domestic sheep breeds evaluated by microsatellite markers [71]. Large frequency of significant LD ( $D'$ ) was observed for syntenic marker-pairs with the interval less than 60 cM. By using 490 markers in Soay sheep [72], 22% marker pairs had  $D' > 0.2$  as for marker-pair distance <

10 cM. Meadows *et al.* [73] used 28 microsatellites in 555 animals from five sheep populations and obtained a flexible LD among different populations. Small LD ranging from 0 to 5 cM was obtained in five populations. In White Faced Suffolk, Poll Dorset and Macarthur Merino populations, average LD extended up to 20 cM in non-syntenic markers. A strong LD was observed at a marker-pair interval of 30 cM in several populations. The  $D'$  value between non-syntenic marker pairs ranged from 0.266 (Poll Dorset) to 0.322 (Merino × Border Leicester).  $r^2$  decayed faster within Merino and Merino × Border Leicester. LD varied a lot among populations in sheep, similar to the results in pig.

## 4. MECHANISMS FOR MAINTAINING PIG GENOMIC DIVERSITY

### 4.1. Migration and Introgression

A significant differentiation in genetic diversity and LD exists between Asian and European pig populations. One reasonable explanation is that their domestications originated from different ancestors. However, they shared some important haplotypes with each other, which might come from extensive migration between populations and incomplete lineage sorting. Historic and demographic events revealed that Near Eastern pigs were definitely introduced into Europe during the 18 - 19<sup>th</sup> centuries. When European wild boars were domesticated, they rapidly became the predominant lineage within European domestic swine [2]. So the nucleotide diversity in Europe had been heavily influenced by the migration from Asia [74]. There was also a widespread haplotype shared between breeds, but a significant genetic distance existed between European and Asian breeds [1, 17, 34, 35]. Even in a very low initial  $N_e$ , subsequent migration can lead to a high level of polymorphism. While maternal introgression from European domestic pig has no or very little impact on Chinese breeds. That is why some European haplotypes were detected in some local breeds, but no European mtDNA haplotypes were detected [35].

### 4.2. Selection

Efficient artificial selection has a dramatic influence on livestock genome diversity and linkage disequilibrium. This effect depends on the direction, intensity, duration and consistency of selection over time. For a long time, in pigs selection focused on growth and fatness (representing demand for cheap lean meat). The availability of dense marker panels provides the opportunity to consider LD between regions and to search for signatures of selection as a means to identify causative mutations or other useful markers as has been done in several species.

For example, two of the first QTL to be identified in the pig were on chromosomes 4 and 7. These include QTL for growth and fatness traits. However, despite significant efforts the genetic variation underlying these QTL has not been elucidated. The extent of LD for SSC7 was significantly larger compared to SSC4 [21]. Further analysis of LD in these regions found that Duroc and Landrace were under the strong selection pressure for both SSC4 and 7 while the selection effect for Large White and Yorkshire was only on SSC7. Interestingly, a significant difference in LD between SSC18 and SSC3 was also obtained (SSC18 > SSC3) that

seemed to correlate with the numbers of candidate genes under selection, as several QTLs were mapped on SSC18 [17]. More recently Ojeda and colleagues have studied a number of gene regions under the SSC4 fatness QTL (FAT1) [22]. Although, they found significantly higher diversity for the region of the gene FABP4, they were unable to find any clear sign of a selective sweep in any of the breeds tested. This study did provide insight into the exchange of germ-plasm between China and Europe, and the authors concluded that an important part of variability within European breeds is due to introgression from Asian pigs. In this case migration is countering the loss of diversity caused by bottlenecks and artificial selection, with this mixing making it potentially harder to identify selective sweeps in pigs.

Modern breeding practices and best linear unbiased prediction (BLUP) selection starting in the last century led to a rapid increase of genetic gain and large LD in pig breeds. This can cause inbreeding that increases LD but reduces genetic diversity. Taking Meishan for example, in order to increase the population size and reproductive performance, an inbred line of Meishan was developed in the late 1980s in a selection scheme. This may explain the lower genetic variability observed within Meishan [19] and that the level of LD was higher [17] compared to those in other Chinese populations.

Selection reduces genetic diversity on the next generation but enhances LD. Selection on one locus or multiple loci, will increase LD between the neighbouring loci and the selected locus. As a result of intensive artificial selection, livestock have a larger LD than human populations. In addition, selection can cause LD between unlinked loci that contribute to phenotypes undergoing selection [18]. This could also affect the structure of genomic diversity in livestock.

### 4.3. Drift

Domestication in livestock essentially results in two processes: the decline in  $N_e$  (bottleneck effects) and directional artificial selection. The reduction in  $N_e$  can lead to lower genetic diversity. LD initially generated by genetic drift gradually reduces with time. The genetic drift effects are substantial in historic pig breeds. The development of European and Chinese pig breeds has been different and the study of these differences can provide additional insights on diversity and these processes. It appears that the differences described to date are the result of differences in the number of ancestral stocks as well as the differences in modern breeding systems. Selection is likely to have been stronger in Europe and N America than in Asia. European breeds had smaller effective population sizes, resulting in stronger LD [17]. American Landrace and Berkshire had smaller  $N_e$  than Hampshire and Duroc breeds, producing different LDs for the same marker-intervals [21, 59]. The joint effects of artificial selection and genetic drift produce different patterns of structured genomic diversity.

## 5. CONCLUSION

In this review, we summarized genomic diversity within and between sites in pig (*Sus scrofa*) and compared them with human and other livestock. It is concluded that genetic diversity is smaller in pig than in human and other livestock.

A relatively larger genetic variation within populations ( $F_{IS}$ ) exists in pig compared with cattle and goat. A majority of genetic variation in pig was distributed within rather than between populations, similar to human and other livestock. A large genetic distance exists between Asian and European populations. Asian breeds harbour much larger variability than European breeds. For the different genetic diversity structures, the inter-site associations described by LD were much stronger in pig (1 - 3 cM) than in human (~ 30 kb) and cattle (~ 100 kb), but less than in sheep (~ 10 cM). LD varied largely among different populations in pig and sheep, but a little in cattle. This review confirms that, strong breeding and selection programs have occurred in pig for a relatively long time, which resulted in the relatively lower heterozygosity and higher  $F_{IS}$  and larger LD, compared with human and cattle. This also implies that a smaller effective population size and less informative markers would be needed for pig whole genome association studies (at least for European populations). To some extent, this difference might arise from the joint effects of migration, selection, and drift during the process of pig domestication.

## ACKNOWLEDGEMENTS

We appreciate the valuable comments of two anonymous reviewers on an earlier version of this article. The work was supported by grants from the Alberta Livestock and Meat Agency and the Natural Sciences and Engineering Research Council of Canada.

## REFERENCES

- [1] Kim, K. I.; Lee, J. H.; Li, K.; Zhang, Y. P.; Lee, S. S.; Gongora, J.; Moran, C. Phylogenetic relationships of Asian and European pig breeds determined by mitochondrial DNA D-loop sequence polymorphism. *Anim. Genet.*, **2002**, *33*(1), 19-25.
- [2] Larson, G.; Dobney, K.; Albarella, U.; Fang, M.; Matisoo-Smith, E.; Robins, J.; Lowden, S.; Finlayson, H.; Brand, T.; Willerslev, E.; Rowley-Conwy, P.; Andersson, L.; Cooper, A. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science*, **2005**, *307*(5715), 1618.
- [3] Giuffra, E.; Kijas, J. M. H.; Amarger, V.; Carlborg, Ö.; Jeon, J. T.; Andersson, L. The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics*, **2000**, *154*(4), 1785-1791.
- [4] Plastow, G.; Siggins, K.; McLaren, D. A genetic case study. Engineering biotechnologies for practical application to livestock. *IEEE Potentials*, **1994**, *13*, 22-25.
- [5] Evans, G. J.; Giuffra, E.; Sanchez, A.; Kerje, S.; Davalos, G.; Vidal, O.; Illán, S.; Noguera, J. L.; Varona, L.; Velander, I.; Southwood, O. I.; de Koning, D. J.; Haley, C. S.; Plastow, G. S.; Andersson, L. Identification of quantitative trait loci for production traits in commercial pig populations. *Genetics*, **2003**, *164*(2), 621-627.
- [6] Hu, Z. L.; Park, C. A.; Fritz, E. R.; Reecy, J. M. QTLdb: A Comprehensive Database Tool Building Bridges between Genotypes and Phenotypes. *Invited Lecture with full paper published electronically on The 9th World Congress on Genetics Applied to Livestock Production. Leipzig, Germany*, **2010**, 1-6, 2010.
- [7] Wei, W. H.; Skinner, T. M.; Anderson, J. A.; Southwood, O. I.; Plastow, G.; Archibald, A. L.; Haley, C. S. Mapping QTL in the porcine MHC region affecting fatness and growth traits in a Meishan/Large White composite population. *Anim. Gen.*, **2010**, DOI: 10.1111/j.1365-2052.2010.02062.x
- [8] Rothschild, M.F.; Newman, S. Eds. Intellectual Property Rights in Animal Breeding and Genetics. *CABI Publishing*, **2002**, 179-196.
- [9] Rothschild, M.; Jacobson, C.; Vaske, D.; Tuggle, C.; Wang, L.; Short, T.; Eckardt, G.; Sasaki, S.; Vincent, A.; McLaren, D.; Southwood, O.; van der Steen, H.; Mileham, A.; Plastow, G. The

- estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*(1), 201-205.
- [10] Short, T. H.; Rothschild, M. F.; Southwood, O. I.; McLaren, D. G.; de Vries, A.; van der Steen, H.; Eckardt, G. R.; Tuggle, C. K.; Helm, J.; Vaske, D. A.; Mileham, A. J.; Plastow, G. S. Effect of estrogen receptor locus on reproduction and production traits in four commercial lines of pigs. *J. Anim. Sci.*, **1997**, *75*(12), 3138-3142.
- [11] van der Steen, H. A. M.; Prall, G. F. W.; Plastow, G. S. Application of genomics to the pork industry. *J. Anim. Sci.*, **2005**, *83*(13), E1-E8.
- [12] de Roos, A. P. W.; Hayes, B. J.; Spelman, R. J.; Godd, M. E. Linkage disequilibrium and persistence of phase in Holstein Friesian, Jersey and Angus cattle. *Genetics*, **2008**, *179*(3), 1503-1512.
- [13] McKay, S. D.; Schnabel, R. D.; Murdock, B. M.; Matukumalli, L. K.; Aerts, J.; Coppieters, W.; Crews, D.; Neto, E. D.; Gill, C. A.; Gao, C.; Mannen, H.; Stothard, P.; Wang, Z.; van Tassell, C. P.; Williams, J. L.; Taylor, J. F.; Moore, S. S. Whole genome linkage disequilibrium maps in cattle. *BMC Genet.*, **2007**, *8*, 74.
- [14] Biswas, S.; Scheinfeldt, L. B.; Akey, J. M. Genome-wide Insights into the Patterns and Determinants of Fine-Scale Population Structure in Humans. *Am. J. Hum. Genet.*, **2009**, *84*(5), 641-650.
- [15] Qanbari, S.; Pimentel, E. C. G.; Tetens, J.; Thaller, G.; Lichtner, P.; Sharifi, A. R.; Simianer, H. The pattern of linkage disequilibrium in German Holstein cattle. *Anim. Genet.*, **2009**, *41*(4), 346-356.
- [16] Upkin, E.; Straus, K.; Stein, T. R.; Bagnato, A.; Schiavini, F.; Fontanesi, L.; Russo, V.; Medugorac, I.; Foerster, M.; Sölkner, J.; Dolezal, M.; Medrano, J. F.; Friedmann, A.; Soller, M. Extensive long-range and nonsyntenic linkage disequilibrium in livestock populations: deconstruction of a conundrum. *Genetics*, **2009**, *181*(2), 691-699.
- [17] Amaral, A. J.; Megens, H.; Crooijmans, R. P. M. A.; Heuven, H. C. M.; Groenen, M. A. M. Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics*, **2008**, *179*(1), 569-579.
- [18] Du, F. X.; Clutter, A. C.; Lohuis, M. M. Characterizing Linkage Disequilibrium in Pig Populations. *Int. J. Biol. Sci.*, **2007**, *3*(3), 166.
- [19] Fan, B.; Wang, Z. G.; Li, Y. J.; Zhao, X. L.; Liu, B.; Zhao, S. H.; Yu, M.; Li, M. H.; Chen, S. L.; Xiong, T. A.; Li, K. Genetic variation analysis within and among Chinese indigenous swine populations using microsatellite markers. *Anim. Genet.*, **2002**, *33*(6), 422-427.
- [20] Luetkemeier, E. S.; Sodhi, M.; Schook, L. B.; Malhi, R. S. Multiple Asian pig origins revealed through genomic analyses. *Mol. Phylogenet. Evol.*, **2010**, *54*(3), 680-686.
- [21] Nsengimana, J.; Baret, P.; Haley, C. S.; Visscher, P. M. Linkage disequilibrium in the domesticated pig. *Genetics*, **2004**, *166*(3), 1395-1404.
- [22] Ojeda, A.; Ramos-Onsins, S. E.; Marletta, D.; Huang, L. S.; Folch, J. M.; Pérez-Enciso, M. Evolutionary study of a potential selection target region in the pig. *Hereditas*, **2011**, *106*, 330-338.
- [23] Ramos, A. M.; Crooijmans, R. P. M. A.; Affara, N. A.; Amaral, A. J.; Archibald, A. L.; Beever, J. E.; Bendixen, C.; Churcher, C.; Clark, R.; Dehais, P.; Hansen, M. S.; Hedegaard, J.; Hu, Z. L.; Kerstens, H. H.; Law, A. S.; Megens, H. J.; Milan, D.; Nonneman, D. J.; Rohrer, G. A.; Rothschild, M. F.; Smith, T. P. L.; Schnabel, R. D.; Van Tassell, C. P.; Taylor, J. F.; Wiedmann, R. T.; Schook, L. B.; Groenen, M. A. M. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS One*, **2009**, *4*(8), e6524.
- [24] Hu, X. S.; Yeh, F. C.; Wang, Z. Structural genomics: correlation blocks, population structure, and genome architecture. *Curr. Genomics*, **2011**, *12*(1), 55-70.
- [25] Laval, G.; Iannuccelli, N.; Legault, C.; Milan, D.; Groenen, M. A. M.; Giuffra, E.; Andersson, L.; Nissen, P. H.; Jørgensen, C. B.; Beeckmann, P.; Geldermann, H.; Foulley, J. L.; Chevalet, C.; Ollivier, L. Genetic diversity of eleven European pig breeds. *Genet. Sel. Evol.*, **2000**, *32*(2), 187-203.
- [26] Yang, S. L.; Wang, Z. G.; Liu, B.; Zhang, G. X.; Zhao, S. H.; Yu, M.; Fan, B.; Li, M. H.; Xiong, T. A.; Li, K. Genetic variation and relationships of eighteen Chinese indigenous pig breeds. *Genet. Sel. Evol.*, **2003**, *35*(6), 657-671.
- [27] The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature*, **2010**, *467*(7319), 1063-1073.
- [28] Fahrenkrug, S. C.; Freking, B. A.; Smith, T. P.; Rohrer, G. A.; Keele, J. W. Single nucleotide polymorphism (SNP) discovery in porcine expressed genes. *Anim. Genet.*, **2002**, *33*(3), 186-195.
- [29] Jungerius, B. J.; Rattink, A. P.; Crooijmans, R. P.; van der Poel, J. J.; van Oost, B. A.; te Pas, M. F.; Groenen, M. A. Development of a single nucleotide polymorphism map of porcine chromosome 2. *Anim. Genet.*, **2003**, *34*(6), 429-43.
- [30] Kerstens, H. H. D.; Kollers, S.; Kommadath, A.; Rosario, M.; Dibbits, B.; Kinders, S. M.; Crooijmans, R. P.; Groene, M. A. M. Mining for Single Nucleotide Polymorphisms in Pig genome sequence data. *BMC Genomics*, **2009**, *10*, 4.
- [31] Wang, D. G.; Fan, J. B.; Siao, C. J.; Berno, A.; Young, P.; Sapolsky, R.; Ghandour, G.; Perkins, N.; Winchester, E.; Spencer, J.; Kruglyak, L.; Stein, L.; Hsie, L.; Topaloglou, T.; Hubbell, E.; Robinson, E.; Mittmann, M.; Morris, M. S.; Shen, N.; Kilburn, D.; Rioux, J.; Nusbaum, C.; Rozen, S.; Hudson, T. J.; Lipshutz, R.; Chee, M.; Lander, E. S. Large-scale identification, mapping and genotyping of single nucleotide polymorphisms in the human genome. *Science*, **1998**, *280*(5366), 1077-1082.
- [32] Archibald, A. L.; Bolund, L.; Churcher, C.; Fredholm, M.; Groenen, M. A. M.; Harlizius, B.; Lee, K. T.; Milan, D.; Rogers, J.; Rothschild, M. F.; Uenishi, H.; Wang, J.; Schook, L. B. Pig genome sequence - analysis and publication strategy. *BMC Genomics*, **2010**, *11*, 438.
- [33] Martinez, A. M.; Delgado, J. V.; Rodero, A.; Vega-Pla, J. L. Genetic structure of the Iberian pig breed using microsatellites. *Anim. Genet.*, **2000**, *31*(5), 295-301.
- [34] Huang, Y. F.; Shi, X. W.; Zhang, Y. P. Mitochondrial Genetic Variation in Chinese Pigs and Wild Boars. *Biochem. Genet.*, **1999**, *37*(11-12), 335-343.
- [35] Fang, M.; Andersson, L. Mitochondrial diversity in European and Chinese pigs is consistent with population expansions that occurred prior to domestication. *Proc. R. Soc. B*, **2006**, *273*(1595), 1803-1810.
- [36] Carza, J. C.; Slatkin, M.; Freimer, N. B. Microsatellite allele frequencies in Humans and Chimpanzees, with implications for constraints on allele size. *Mol. Biol. Evol.*, **1995**, *12*(4), 594-603.
- [37] Thalmann, O.; Fischer, A.; Lankester, F.; Pääbo, S.; Vigilant, L. The complex evolutionary history of gorillas: insights from genomic data. *Mol. Biol. Evol.*, **2007**, *24*(1), 146-158.
- [38] The Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, **2005**, *437*(7055), 69-87.
- [39] Auton, A.; Bryc, K.; Boyko, A.; Lohmueller, K.; Novembre, J.; Reynolds, A.; Indap, A.; Wright, M. H.; Degenhardt, J.; Gutenkunst, R.; King, K. S.; Nelson, M. R.; Bustamante, C. D. Global distribution of genomic diversity underscores rich complex history of continental human populations. *Genome Res.*, **2009**, *19*(5), 795-803.
- [40] Barreiro, L. B.; Laval, G.; Quach, H.; Patin, E.; Quintana-Murci, L. Natural selection has driven population differentiation in modern humans. *Nat. Genet.*, **2008**, *40*(3), 340-345.
- [41] International Hap Map Consortium. A haplotype map of the human genome. *Nature*, **2005**, *437*(7063), 1299-1320.
- [42] Weir, B. S.; Cardon, L. R.; Anderson, A. D.; Nielsen, D. M.; Hill, W. G. Measures of human population structure show heterogeneity among genomic regions. *Genome Res.*, **2005**, *15*(11), 1468-1476.
- [43] Xing, J.; Watkins, W. S.; Witherspoon, D. J.; Zhang, Y.; Guthery, S. L.; Thara, R.; Mowry, B. J.; Bulayeva, K.; Weiss, R. B.; Jorde, L. B. Fine-scaled human genetic structure revealed by SNP microarrays. *Genome Res.*, **2009**, *19*(5), 815-825.
- [44] Redon, R.; Ishikawa, S.; Fitch, K. R.; Feuk, L.; Perry, G. H.; Andrews, T. D.; Fiegler, H.; Shaperro, M. H.; Carson, A. R.; Chen, W.; Cho, E. K.; Dallaire, S.; Freeman, J. L.; González, J. R.; Gratacòs, M.; Huang, J.; Kalaitzopoulos, D.; Komura, D.; MacDonald, J. R.; Marshall, C. R.; Mei, R.; Montgomery, L.; Nishimura, K.; Okamura, K.; Shen, F.; Somerville, M. J.; Tchinda, J.; Valsesia, A.; Woodwark, C.; Yang, F.; Zhang, J.; Zerjal, T.; Zhang, J.; Armenoglou, L.; Conrad, D. F.; Estivill, X.; Tyler-Smith, C.; Carter, N. P.; Aburatani, H.; Lee, C.; Jones, K. W.; Scherer, S. W.; Hurles, M. E. Global variation in copy number in the human genome. *Nature*, **2007**, *444*(7118), 444-454.
- [45] Bruno-de-Sousa, C.; Martinez, A. M.; Ginja, C.; Santos-Silva, F.; Carolino, M. I.; Delgado, J. V.; Gama, L. T. Genetic diversity and population structure in Portuguese goat breeds. *Livest. Sci.*, **2010**, doi:10.1016/j.livsci.2010.06.159.



- [46] Dixit, S. P.; Verma, N. K.; Aggarwal, R. A. K.; Kumar, S.; Chander, R.; Vyas, M. K.; Singh, K. P. Genetic structure and differentiation of three Indian goat breeds. *Asian-Aust. J. Anim. Sci.*, **2009**, *22*, 1234-1240.
- [47] Serrano, M.; Calvo, J. H.; Martinez, M.; Marcos-Carcavilla, A.; Cuevas, J.; Gonzallez, C.; Jurado, J. J.; Tejada, P. D. Microsatellite based genetic diversity and population structure of the endangered Spanish Guadarrama goat breed. *BMC Genet.*, **2009**, *10*, 61.
- [48] Agha, S.H.; Pilla, F.; Galal, S.; Shaat, I.; Andrea, M. D.; Reale, S.; Abdelsalam, A. Z. A.; Li, M. H. Genetic diversity in Egyptian and Italian goat breeds measured with microsatellite polymorphism. *J. Anim. Breed. Genet.*, **2008**, *125*(3), 194-200.
- [49] Chenyambuga, S. W.; Hanotte, O.; Hirbo, J.; Watts, P. C.; Kemp, S. J.; Kifaro, G. C.; Gwakisa, P. S.; Petersen, P. H.; Rege, J. E. O. Genetic characterization of indigenous goats of Sub-Saharan Africa using microsatellite DNA markers. *Asian-Aust. J. Anim. Sci.*, **2004**, *17*(4), 445-452.
- [50] Zhou, G. L.; Jin, H. G.; Zhu, Q.; Guo, S. L.; Wu, Y. H. Genetic diversity analysis of five cattle breeds native to China using microsatellites. *J. Genet.*, **2005**, *84*(1), 77-80.
- [51] Canon, J.; Alexandrino, P.; Bessa, I.; Carleos, C.; Carretero, Y.; Dunner, S.; Ferran, N.; Garcia, D.; Jordana, J.; Laloe, D.; Pereira, A.; Sanchez, A.; Moazami-Goudarzi, K. Genetic diversity measures of local European beef cattle breeds for conservation purposes. *Genet. Sel. Evol.*, **2001**, *33*(3), 311-332.
- [52] Kantanen, J.; Olsaker, I.; Holm, L. E.; Lien, S.; Vilkki, J.; Brusgaard, K.; Eythorsdottir, E.; Danell, B.; Adalsteinsson, S. Genetic diversity and population structure of 20 north European cattle breeds. *J. Hered.*, **2000**, *91*(6), 446-457.
- [53] Egitto, A. A.; Paiva, S. R.; Albuquerque, M. S. M.; Mariante, A. S.; Almeida, L. D.; Castro, S. R.; Grattapaglia, D. Microsatellite based genetic diversity and relationships among ten Creole and commercial cattle breeds raised in Brazil. *BMC Genet.*, **2007**, *8*, 83.
- [54] Rehman, M. S.; Khan, M. S. Genetic diversity of haryana and hissar cattle from Pakistan using microsatellite analysis. *Pakistan Vet. J.*, **2009**, *29*(2), 67-71.
- [55] Harmegnies, N.; Farnir, F.; Davin, F.; Buys, N.; Georges, M.; Coppieters, W. Measuring the extent of linkage disequilibrium in commercial pig populations. *Anim. Genet.*, **2006**, *37*(3), 225-231.
- [56] Uimari, P.; Tapio, M. Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. *J. Anim. Sci.*, **2010**, doi:10.2527/jas.2010-3249.
- [57] Jungerius, B. J.; Gu, J.; Crooijmans, R. P. M. A.; Poel, J. J.; Groenen, M. A. M.; van Oost, B. A.; te Pas, M. F. Estimation of the extent of linkage disequilibrium in seven regions of the porcine genome. *Anim. Biotechnol.*, **2005**, *16*(1), 41-54.
- [58] Megens, H. J.; Crooijmans, R. P. M. A.; Heuven, H. C. M.; Groenen, M. A. M. Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics*, **2008**, *179*(1), 569-579.
- [59] Welsh, C. S.; Stewart, T. S.; Schwab, C.; Blackburn, H. D. Pedigree analysis of 5 swine breeds in the United States and the implications for genetic conservation. *J. Anim. Sci.*, **2010**, *88*(5), 1610-1618.
- [60] Dunning, A. M.; Durocher, F.; Healey, C. S.; Teare, M. D.; McBride, S. E.; Carlomagno, F.; Xu, C. F.; Dawson, E.; Rhodes, S.; Ueda, S.; Lai, E.; Luben, R. N.; van Rensburg, E. J.; Mannermaa, A.; Kataja, V.; Rennart, G.; Dunham, I.; Purvis, I.; Easton, D.; Ponder, B. A. J. The extent of linkage disequilibrium in four populations with distinct demographic histories. *Am. J. Hum. Genet.*, **2000**, *67*(6), 1544-1554.
- [61] Kruglyak, L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.*, **1999**, *22*(2), 139-144.
- [62] Abecasis, G. R.; Noguchi, E.; Heinzmann, A.; Traherne, J. A.; Bhattacharyya, S.; Leaves, N. I.; Anderson, G. G.; Zhang, Y.; Lench, N. J.; Carey, A.; Cardon, L. R.; Moffatt, M. F.; Cookson, W. O. C. Extent and distribution of linkage disequilibrium in three genomic regions. *Am. J. Hum. Genet.*, **2001**, *68*(1), 191-197.
- [63] Taillon-Miller, P.; Bauer-Sardiña, I.; Saccone, N. L.; Putzel, J.; Laitinen, T.; Cao, A.; Kere, J.; Pilia, G.; Rice, J. P.; Kwok, P. Y. Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. *Nat. Genet.*, **2000**, *25*(3), 324-328.
- [64] Ardlie, K. G.; Kruglyak, L.; Seielstad, M. Patterns of linkage disequilibrium in the human genome. *Nat. Rev. Genet.*, **2002**, *3*(4), 299-309.
- [65] Reich, D. E.; Cargill, M.; Bolk, S.; Ireland, J.; Sabeti, P. C.; Richter, D. J.; Lavery, T.; Kouyoumjian, R.; Farhadian, S. F.; Ward, R.; Lander, E. S. Linkage disequilibrium in the human genome. *Nature*, **2001**, *411*(6834), 199-204.
- [66] Tenesa, A.; Navarro, P.; Hayes, B. J.; Duffy, D. L.; Clarke, G. M.; Goddard, M. E.; Visscher, P. M. Recent human effective population size estimated from linkage disequilibrium. *Genome Res.*, **2007**, *17*(4), 520-526.
- [67] Flury, C.; Tapio, M.; Sonstegard, T.; Drögemüller, C.; Leeb, T.; Simianer, H.; Hanotte, O.; Rieder, S. Effective population size of an indigenous Swiss cattle breed estimated from linkage disequilibrium. *J. Anim. Breed. Genet.*, **2010**, *127*(5), 339-347.
- [68] Khatkar, M. S.; Nicholas, F. W.; Collins, A. Q.; Zenger, K. R.; Cavanagh, J. A. L.; Barris, W.; Schnabel, R. D.; Taylor, J. F.; Raadma, H. W. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. *BMC Genomics*, **2008**, *9*, 187.
- [69] Sargolzaei, M.; Schenkel, F. S.; Jansen, G. B.; Schaeffer, L. R. Extent of linkage disequilibrium in Holstein cattle in North America. *J. Dairy Sci.*, **2008**, *91*(5), 2106-2117.
- [70] Kim, E. S.; Kirkpatrick, B. W. Linkage disequilibrium in the North American Holstein population. *Anim. Genet.*, **2009**, *40*(3), 279-88.
- [71] McRae, A. F.; McEwan, J. C.; Dodds, K. G.; Wilson, T.; Crawford, A. M.; Slate, J. Linkage disequilibrium in domestic sheep. *Genetics*, **2002**, *160*(3), 1113-1122.
- [72] McRae, A. F.; Pemberton, J. M.; Visscher, P. M. Modeling linkage disequilibrium in natural populations: the example of the soay sheep population of St. Kilda, Scotland. *Genetics*, **2005**, *171*(1), 251-258.
- [73] Meadows, J. R. S.; Chan, E. K.; Kijas, J. W. Linkage disequilibrium compared between five populations of domestic sheep. *BMC Genet.*, **2008**, *9*, 61.
- [74] Ramirez, O.; Ojeda, A.; Tomas, A.; Gallardo, D.; Huang, L. S.; Folch, J. M.; Clop, A.; Sánchez, A.; Badaoui, B.; Hanotte, O.; Galman-Omitogun, O.; Makuza, S. M.; Soto, H.; Cadillo, J.; Kelly, L.; Cho, I. C.; Yeghoyan, S.; Pérez-Enciso, M.; Amills, M. Integrating Y-chromosome, mitochondrial and autosomal data to analyse the origin of pig breeds. *Mol. Biol. Evol.*, **2009**, *26*(9), 2061-2072.
- [75] Xing, J.; Watkins, W. S.; Shlien, A.; Walker, E.; Huff, C. D.; Witherspoon, D. J.; Zhang, Y.; Simonson, T. S.; Weiss, R. B.; Schifman, J. D.; Malkin, D.; Woodward, S. R.; Jorde, L. B. Toward a more uniform sampling of human genetic diversity: A survey of worldwide populations by high-density genotyping. *Genomics*, **2010**, *96*(4), 199-210.