Molecular Phylogenetics and Evolution 54 (2010) 680–686

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Multiple Asian pig origins revealed through genomic analyses

Erin S. Luetkemeier [a], Monika Sodhi [a], Lawrence B. Schook [a,c], Ripan S. Malhi [b,c,*]

[a] Department of Animal Sciences, University of Illinois, 374 Edward R. Madigan Lab, 1201 West Gregory Drive, Urbana, IL 61801, USA
[b] Department of Anthropology, University of Illinois, 109 Davenport Hall, 607 S. Mathews Ave., Urbana, IL 61801, USA
[c] Institute for Genomic Biology, University of Illinois, Urbana, IL 61801, USA

ARTICLE INFO

Article history:
Received 14 October 2008
Revised 2 November 2009
Accepted 3 November 2009
Available online 10 November 2009

Keywords:
European domestic pig populations
Asian domestic pig populations
Microsatellite markers
Population structure
Wild boars

ABSTRACT

Previous mitochondrial DNA (mtDNA) studies have suggested that European and Asian pig populations were derived through multiple domestication events. We investigated whether domestic pig populations were derived from distinct ancestors within their respective regions, using eight domestic breeds (five European and three Asian), and also European and Asian wild boar populations. Genomic analyses utilized 21 microsatellite markers (MS) selected for their distribution across the pig genome in addition to the mtDNA D-loop region. The number of alleles per MS loci ranged from 8 (Sw2008) to 16 (S0097 and S0218). Few significant departures from Hardy–Weinberg equilibrium were detected, suggesting the absence of heterozygote deficiencies. Analyses within populations revealed observed mean heterozygosity from 0.48 (Erhualian) to 0.68 (Dutch WB) and an expected mean heterozygosity from 0.53 (Hampshire) to 0.80 (Japanese WB) with effective alleles ranging from 2.28 (Hampshire) to 3.74 (French WB). Wild boar populations demonstrated a higher level of heterozygosity than domestic breeds. Genetic differentiation estimated by fixation indices (F_{ST}) ranged from 0.021 (Yorkshire and Duroc) to 0.410 (Meishan and Hampshire) and was consistent with previous mtDNA analysis. Both phylogenetic and principal component analyses revealed a distinct separation of European and Asian derived populations with tight clustering of the European domestic breeds. Conversely, the use of both MS and mtDNA clarified that the Asian populations were comprised of three groups, one represented by Erhualian and Meishan breed, the second represented by Lanyu pigs and the third represented by the Asian wild boars. The current findings support the hypothesis that Asian domestic populations were derived from multiple Asian ancestral origins whereas the European domestic populations represent a single ancestral European lineage.

Published by Elsevier Inc.

1. Introduction

The wild boar is found distributed throughout Eurasia and northern Africa with close to 25 classified subspecies (Clutton-Brock, 1987). The pig was proposed to be domesticated independently from subspecies of the European and Asian wild boar populations (Okumura et al., 2001). Mitochondrial DNA (mtDNA) analyses have provided evidence for multiple independent domestication events of the European and Asian subspecies of wild boar (Kurosawa et al., 1984; Oishi et al., 1993; Giuffra et al., 2000). Within Europe and Asia, Larson et al. (2005) identified six major origin clusters suggesting that domestication occurred from genetically distinct wild boar subpopulations. Additional studies demonstrated the influence of human interaction on the structure of domestic pig populations through the introgression of Asian germplasm into European pig breeds over the past 200 years (Larson et al., 2005). Most evolutionary studies to date have utilized mtDNA for analysis. The effective population size of mtDNA is lower than that of nuclear DNA, making it useful for addressing genetic relationships within and between closely related populations and the presence of population bottlenecks (Berggren et al., 2005). However, mtDNA may be a poor indicator of overall genomic diversity because it is a single locus and is an extra-nuclear genetic marker with unique evolutionary dynamics. Also, it is maternally inherited and does not detect male-mediated gene flow, which has been a powerful influence on the evolution of pigs in modern times (Pamilo and Nei, 1988). Conversely, information obtained from multiple polymorphic genomic markers in linkage equilibrium, such as (MS) can provide a genomic approach to revealing evolutionary relationships among pig breeds and wild boar populations. Due to their abundance, genome distribution and the lack of ascertainment bias, MS serve as powerful markers for evaluating genetic diversity. Additionally, MS are proven tools for investigating genetic relatedness (Forbes et al., 1995; Estoup et al., 1996). Although, the best use of neutral markers to estimate the impact of population historic events would be the use of multi-

* Corresponding author. Address: Department of Anthropology, University of Illinois, 109 Davenport Hall, 607 S. Mathews Ave., Urbana, IL 61801, USA. Fax: +1 217 244 3490.
E-mail address: malhi@uiuc.edu (R.S. Malhi).

1055-7903/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.ympev.2009.11.004

ple types of markers in combination (Bruford et al., 2003). The novelty of this study is the use of both mtDNA and MS, where the combination of these markers would avoid inheritance bias since they relay information on maternally and codominantly inherited regions (Yue et al., 1999).

Previous mtDNA analyses demonstrated a divergence of European and Asian pigs that supported the hypothesis of independent origins of domestic pig populations (Fan et al., 2002). Previous MS analyses within European domestic pigs revealed the close clustering of domestic breeds with minor branch separation of Yorkshire and Duroc breeds from Landrace and Meishan breeds (Li and Enfield, 1989). The grouping of Asian domestic pig breeds identified by MS clustering was in accordance with their geographic origin, distribution, body conformation and color (Zhang, 1986) indicating Asian indigenous pig breeds have a close relationship and most likely originated from a common ancestor (Yang et al., 2003). It can be difficult to determine the relationship between European and Asian pigs when different genetic markers were used for the analysis. Hence to further resolve ancestral origins, we performed a comprehensive study involving both domestic and wild boar populations from Europe and Asia to determine the origins and relationships of both geographic locations simultaneously.

2. Materials and methods

2.1. Animals

A total of 149 unrelated European and Asian pigs (no common grandparents) representing four different pig populations (European domestics, Asian domestics, European wild boars and Asian wild boars) were used for MS and mtDNA analyses (Table 1). Wild boar is a term that refers to *Sus scrofa* that has not been domesticated and is a different subspecies from *S. scrofa domestica*, Wild boar and domestic pigs samples were distinguished by previously approved morphological differences in the animals (Jonsson, 1991). Domestic pig samples included five domestic breeds of European origin (Yorkshire, Hampshire, Duroc, Large White, and Landrace), three domestic breeds of Asian origin (Erhualian, Lanyu and Meishan), three European wild boar subpopulations (Italian, Dutch and French wild boar) and two Asian wild boar subpopulations (Japanese wild boar and Ryukyu wild boar). DNA from the domestic populations was obtained from different suppliers throughout the United States, Taiwan and Beijing in an attempt to decrease the likelihood of common origins. DNA from wild boar samples were obtained from Sardinia (Italy), Yamanashi, Ooita, Ehime and Kagoskima Prefectures (Japan), Deux Serves (France) and Veluwezoom and Meinweg (The Netherlands). Additionally, within each population, unrelated animals (no shared grandparents) were used to decrease the probability of genetic relatedness thus eliminating false phylogenetic relationships due to familial relationships.

2.2. Microsatellite genotyping

The International Society of Animal Genetics (ISAG) and the Food and Agriculture Organization (FAO) have recommended a set of 30 MS (http://www2.toulouse.inra.fr/lgc/pig/panel/panel2004.htm) for evaluating pig genetic diversity as part of their global strategy for the management of farm animal genetic resources (Hammond and Leitch, 1998). The 30 MS recommended by FAO-ISAG for diversity studies in swine were used (see Table 2). The MS markers were previously identified due to high polymorphism, absence of null alleles and genome distribution. Primers were synthesized by Integrated DNA technologies (IDT, Coralville, IA). Reactions were designed to use forward primers tagged with the universal M13 forward primer and the standard reverse primers. The M13 universal primer was labeled at the 5′ end with either of four different Beckman Coulter dyes. PCR reactions were carried out using 45 ng DNA in 20 μl reaction mixture containing 2 μl of 10× Buffer, 1 μl of 4 mM dNTPs (Invitrogen, Carlsbad, CA), 0.1 μl of HotStarTaq Polymerase (Qiagen, Valencia, CA), 1 μM tagged forward primer, and 10 μM reverse and labeled M13 primer. PCR conditions had an initial denaturing temperature of 95 °C for 5 min followed by 35 cycles of a three step process of 30 s at 94 °C, 30 s at 58 °C, and 1 min at 72 °C followed by a final step of 45 min at 72 °C. PCR products were analyzed by electrophoresis in a 3% agarose 1× TBE gel and visualized by ethidium bromide staining and UV light. Positive PCR products

Table 1
Average within-breed genetic variation (21 microsatellite loci).

Source	n	N_a	N_e	H_e	H_o
Yorkshire	12	4.86	3.20	0.69	0.65
Hampshire	12	3.57	2.29	0.53	0.54
Duroc	8	3.52	2.74	0.65	0.59
Large White	11	3.43	2.44	0.57	0.54
Landrace	12	4.67	3.09	0.67	0.65
All European domestics	55	4.01	2.75	0.62	0.59
Erhualian	5	2.95	2.39	0.60	0.48
Lanyu	12	3.95	2.49	0.60	0.54
Meishan	12	4.05	2.64	0.58	0.54
All Asian domestics	29	3.65	2.51	0.59	0.52
Italian wild boar	14	5.00	3.00	0.66	0.65
French wild boar	16	6.19	3.74	0.76	0.61
Dutch wild boar	20	5.33	3.33	0.68	0.68
All European wild boars	50	5.51	3.36	0.70	0.65
Japanese wild boar	8	4.72	3.22	0.72	0.65
Ryukyu wild boar	7	4.95	3.45	0.72	0.55
All Asian wild boars	15	6.76	4.60	0.79	0.60
All	139	4.09	3.43	0.81	0.61

Abbreviations: (n) number of individuals included, (N_a) number of alleles, (N_e) number of effective alleles, (H_e) expected heterozygosity, (H_o) observed heterozygosity.

Table 2
Genetic variation of microsatellite loci.

	Chromosome	Marker	N_a	H_e	H_o	F_{IS}	F_{ST}
1	1q	SO155*	9	0.793	0.550	0.097	0.223
2	2q	S0226*	11	0.846	0.508	0.126	0.374
3	3p	Sw72	11	0.785	0.595	0.018	0.296
4	3q	S0002	14	0.838	0.618	−0.091	0.321
5	4	S0097*	16	0.844	0.536	0.103	0.257
6	5q	S005*	14	0.827	0.543	0.107	0.233
7	6q	S0228	15	0.843	0.636	−0.028	0.246
8	7q	Sw632	14	0.833	0.689	−0.141	0.260
9	7q	S0101*	10	0.771	0.602	0.113	0.165
10	8	Sw2410	11	0.628	0.461	0.043	0.233
11	8q	S0178	10	0.806	0.792	−0.168	0.181
12	9p	Sw911	11	0.775	0.706	−0.106	0.219
13	10	Sw830	9	0.690	0.533	−0.038	0.272
14	11	Sw2008	8	0.704	0.403	0.020	0.414
15	13	Swr1941	9	0.828	0.653	−0.088	0.277
16	14q	SW857	11	0.808	0.568	0.143	0.197
17	15q	S0355*	9	0.657	0.288	0.194	0.421
18	15q	Sw936	11	0.833	0.550	0.100	0.308
19	16q	S0026	10	0.764	0.630	−0.112	0.232
20	17q	Sw24	13	0.757	0.543	−0.076	0.295
21	Xq	S0218*	16	0.782	0.662	0.000	0.212
		Mean*	11.62	0.812	0.610	0.015	0.261

Abbreviations: (N_a) number of alleles, (H_e) expected heterozygosity, (H_o) observed heterozygosity, (F_{IS}) within population inbreeding estimate and (F_{ST}) measurement of population differentiation.
* Significant ($P < 0.05$) departure from Hardy–Weinberg equilibrium.

were purified using a Qiagen MinElute 96 UF PCR purification kit (Qiagen, Valencia CA). All products were sequenced using an ABI Prism 3730xl DNA capillary sequencer (Applied Biosystems, Foster, CA). Alleles were scored based on size relative to known DNA size standards and analyzed using the fragment analysis software GeneMarker (SoftGenetics, State College, PA).

2.3. Mitochondrial D-loop region amplification and sequencing

The D-loop region was amplified using the following primers previously published in Larson et al. (2005); L strand (L15387) 5′CTCCGCCATCAGCACCCAAAG3′ and H strand (H16108n) 5′GCACCT TGTTTGGATTRTCG3′. PCR was performed using 20 ng DNA in a 20 μl volume reaction containing 2 μl of 10× buffer, 1 μl of 4 mM dNTPs (Invitrogen, Carlsbad, CA), 0.1 μl of HotStarTaq Polymerase (Qiagen, Valencia, CA), and 0.1 μl of each 100 mM primer. PCR conditions had an initial denaturing temperature of 95 °C for 10 min followed by 35 cycles of a three step process of 30 s at 94 °C, 30 s at 54 °C, and 1 min at 72 °C followed by a final step of 7 min at 72 °C. PCR products were analyzed by electrophoresis in a 1% agarose 1× TBE gel and visualized by ethidium bromide staining and UV light. Positive PCR products were purified using a Qiagen MinElute 96 UF PCR purification kit (Qiagen, Valencia CA). Bidirectional sequencing reactions were performed in 10 μl containing 2 μl ABI BigDye, 2 μl of 5× sequencing buffer, 0.25 μl of primer, and 2 μl purified PCR template. Cycle-sequencing conditions had an initial denaturing temperature of 96 °C for 4 min followed by 35 cycles of a three step process of 15 s at 96 °C, 10 s at 55 °C, and 4 min at 60 °C.

2.4. Data analysis

For each locus and population, different measures of genetic variation, particularly number of alleles (N_a), effective number of alleles (n_e), observed heterozygosity (H_o) and expected heterozygosity (H_e) were estimated using the POPGENE software package (Yeh et al., 1999). Populations experiencing heterozygous deficiency revealed by significant departures from Hardy–Weinberg equilibrium (HWE) ($P < 0.05$) were detected using GENEPOP version 3.1 (Raymond and Rousset, 1995). STRUCTURE2.2 (Pritchard et al., 2000) was used for the population cluster analysis. Parameters were set to run simulations for values of K (number of populations) from 1 to 13, and for each value of K, the Burnin period was set to 50,000 repetitions followed by 500,000 Markov chain Monte Carlo (MCMC) repetitions. The appropriate number of clusters was estimated in two ways: (1) utilization of best fit log likelihood values and (2) performing a ΔK calculation, which is an ad hoc quantity related to the second order rate of change of the log probability of data with respect to the number of clusters (Evanno et al., 2005).

In the assessment of population structures, the ARLEQUIN version 4.0 (Excoffier et al., 2005) computer package was utilized to calculate fixation indices (F_{IS} and F_{ST}) and genetic diversity. F-statistics were used to estimate population substructure within (F_{IS}) and between populations (F_{ST}). F_{ST} estimates were derived from the variance-based method of Weir and Cockerham (1984). The number of genetically diverged groups in a population was estimated by Bayesian inference of the genetic structure (BAPS4) using MS genetic distances to construct a phylogenetic tree (Corander et al., 2004). For mtDNA analysis, Nei's genetic distances and maximum likelihood among populations were calculated using MEGA version 2.0 (Kumar et al., 2001) and used to draw phylogenetic trees by an unweighted pair-group method using an arithmetic mean method (UPGMA) and Neighbor-joining (NJ). The numbers on the branches indicates the percentage occurrence in 1000 bootstrap replicates and the scale indicates genetic distance. The principal components analysis (PCA) was performed using the JMP7 statistical program (SAS institute, Cary, NC).

3. Results

3.1. Allelic and locus variation

The current study analyzed DNA polymorphism in 149 animals. Although each of the 30 MS loci were amplified, only the 21 MS amplicons yielding >4 alleles/loci for all of the populations tested were included in the analyses in accordance with FAO guidelines for population diversity analyses (http://dad.fao.org, Secondary Guidelines MoDAD). Statistical parameters pertaining to genetic variation, such as number of alleles (N_a), observed and expected heterozygosity (H_o and H_e) and F-statistics (F_{IS} and F_{ST}) were calculated for each MS marker (Table 2). A total of 242 alleles were detected for the 21 loci, with an average of 11.62 alleles per locus. The number of alleles ranged from 8 at locus Sw2008 to 16 at locus S0097 and S0218. Each of the MS amplified were polymorphic in all populations examined except for S0101, Sw72 and S0355 that were monomorphic in the Large White, the Erhualian and the Hampshire breeds, respectively. Observed and expected heterozygosity ranged from 0.228 (S0355) to 0.792 (S0178) and from 0.628 (Sw2410) to 0.846 (S0226), respectively.

3.2. Within-population diversity and population structure

The level of diversity within each population was determined using the number of alleles (N_a), the effective number of alleles (N_e), the expected heterozygosity (H_e) and the observed heterozygosity (H_o). The two measures of heterozygosity, (H_e and H_o) are closely related but significant differences between the two suggest that populations have either a heterozygous deficiency or the presence of null alleles. The observed mean MS heterozygosity ranged from 0.48 (Erhualian) to 0.68 (Dutch wild boar), whereas the expected mean MS heterozygosity ranged from 0.53 (Hampshire) to 0.76 (French wild boar), with the number of effective MS alleles ranging from 2.29 (Hampshire) to 3.74 (French wild boar) (Table 1). The wild boar populations revealed a higher level of diversity than the domestic populations when the number of alleles and level of heterozygosity were compared. The number of observed MS alleles in the domestic populations (European: 4.01 and Asian: 3.65) and the level of observed MS heterozygosity (European: 0.594 and Asian: 0.520) revealed that the wild boar populations contained more alleles (European: 5.51 and Asian: 6.76) with a higher level of observed MS heterozygosity (European: 0.65 and Asian: 0.60) than the domestic animals (Table 1).

Assignment of each individual to the correct population was carried out by cluster analysis using a Bayesian approach based on genotypes from multiple loci. The Bayesian approach accounts for HWE or linkage disequilibrium by forming populations that are not in disequilibrium. The appropriate number of populations thus determined was five: (1) a European domestic population; (2) an Asian domestic population (excluding the Taiwanese Lanyu); (3) a Lanyu population; (4) a European wild population; and (5) an Asian wild boar population (Fig. 1). Populations that were determined are in accordance with previous groupings based on the geographical distribution and morphological characteristics (Clutton-Brock, 1987).

3.3. Between-population diversity: implications for regional geographic genetic breakdown

To investigate whether domestication within Europe and Asia were independent events, F_{ST} estimates and a PCA were performed using European and Asian wild boar populations. Domestic populations of the two continents showed the highest degree of genetic differentiation with a F_{ST} value of 0.25 observed between all the

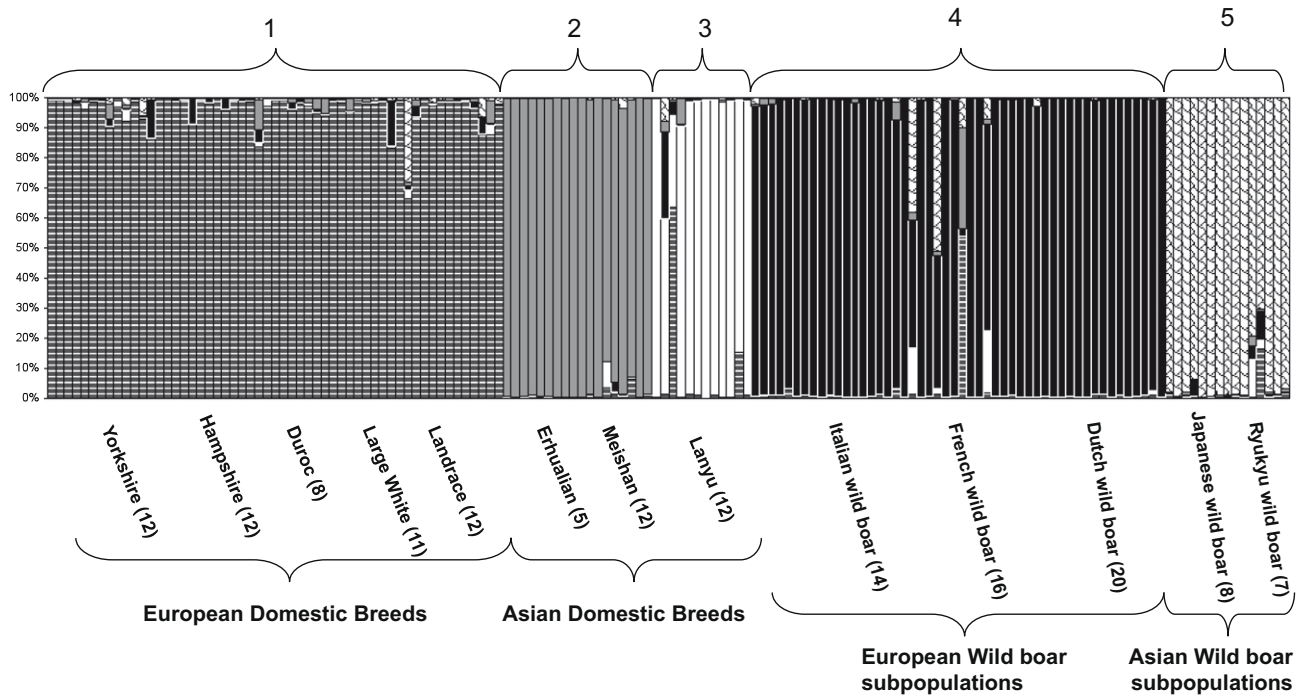

Fig. 1. Population structures defined through bayesian clustering. Clustering was performed under the assumption of $K = 5$. Each bar represents a different individual. Number of individuals per sample group in parentheses.

Asian and European domestic populations. The F_{ST} observed between the Asian wild boars and European wild boars was comparatively low (0.19). The higher F_{ST} values in the domestic populations most likely reflected years of independent breeding within each region following domestication. The smallest pairwise F_{ST} estimate between sample groups (Table 3) was 0.021, between the Duroc and Yorkshire breeds, creating a single cluster. The largest genetic differentiation observed was 0.410 between the Meishan and Hampshire breeds.

Genetic relationships between European and Asian populations were calculated for both MS and mtDNA genotypes using Nei's genetic distances, Bayesian analysis, and maximum likelihood to infer phylogenies by the unweighted pair-group method with arithmetic mean (UPGMA) and Neighbor-joining (NJ) trees. No significant differences between the three methods were demonstrated and so only the UPGMA tree was used for analyses. Both for MS and mtDNA genotypes, the four populations (European domestics, Asian domestics, European wild boars, and Asian wild boars) analyzed were divided into two main branches, one representing the European groups and the other for the Asian groups (Fig. 2 and Suppl. Fig. I). In analysis of the MS Nei's genetic distances phylogenetic tree, within each main branch for European and Asian derived animals, the wild boar populations clustered together at distinct branches. Within the European branch, the domestic pig breeds had smaller genetic distances resulting in closer clustering than the Asian domestic breeds. The only difference identified in the MS and mtDNA phylogenetic trees was the clustering of the Lanyu breed. In the genomic MS UPGMA analysis, the Lanyu breed clustered with the European populations, whereas in the mtDNA analysis it clustered with the Asian populations. The variation in clustering identified by different markers suggests that the Lanyu breed may have maternal Asian origins with European genomic contributions.

To present geometric relationships among the pig populations, a principal component analysis (PCA) was applied using gene frequencies of all loci. The frequencies of all alleles at a single locus were considered to be independent variables. A scatterplot of the score data was examined to visualize the geometric relationships among the pig populations. Consistent with the phylogenetic tree results, the PCA plot revealed a tight clustering for the breeds in

Table 3
Matrix of population differentiation (F_{ST}) between sample groups using microsatellite markers.

	Yorkshire	Hampshire	Duroc	Large White	Landrace	Erhualian	Lanyu	Meishan	Italian WB	France WB	Dutch WB	Japanese WB	Ryukyu
WB													
Yorkshire													
Hampshire	0.20923												
Duroc	0.02128	0.18286											
Large White	0.17676	0.10663	0.15026										
Landrace	0.11806	0.14798	0.08081	0.14515									
Erhualian	0.18399	0.29662	0.18669	0.24327	0.19873								
Lanyu	0.19918	0.28682	0.16872	0.21895	0.2002	0.20559							
Meishan	0.29557	0.4095	0.27751	0.35517	0.31486	0.18202	0.29434						
Italian WB	0.18397	0.2238	0.13674	0.20196	0.18782	0.17957	0.22302	0.31014					
French WB	0.15618	0.18868	0.11597	0.16612	0.11939	0.15895	0.18636	0.26096	0.12639				
Dutch WB	0.18156	0.17676	0.14184	0.17597	0.14153	0.21675	0.23221	0.31418	0.15921	0.07516			
Japanese WB	0.22206	0.31305	0.19167	0.25643	0.2316	0.14557	0.21747	0.2612	0.24654	0.20153	0.23329		
Ryukyu WB	0.20952	0.33995	0.18111	0.30177	0.23608	0.14493	0.23701	0.25693	0.23342	0.1917	0.24998	0.13937	

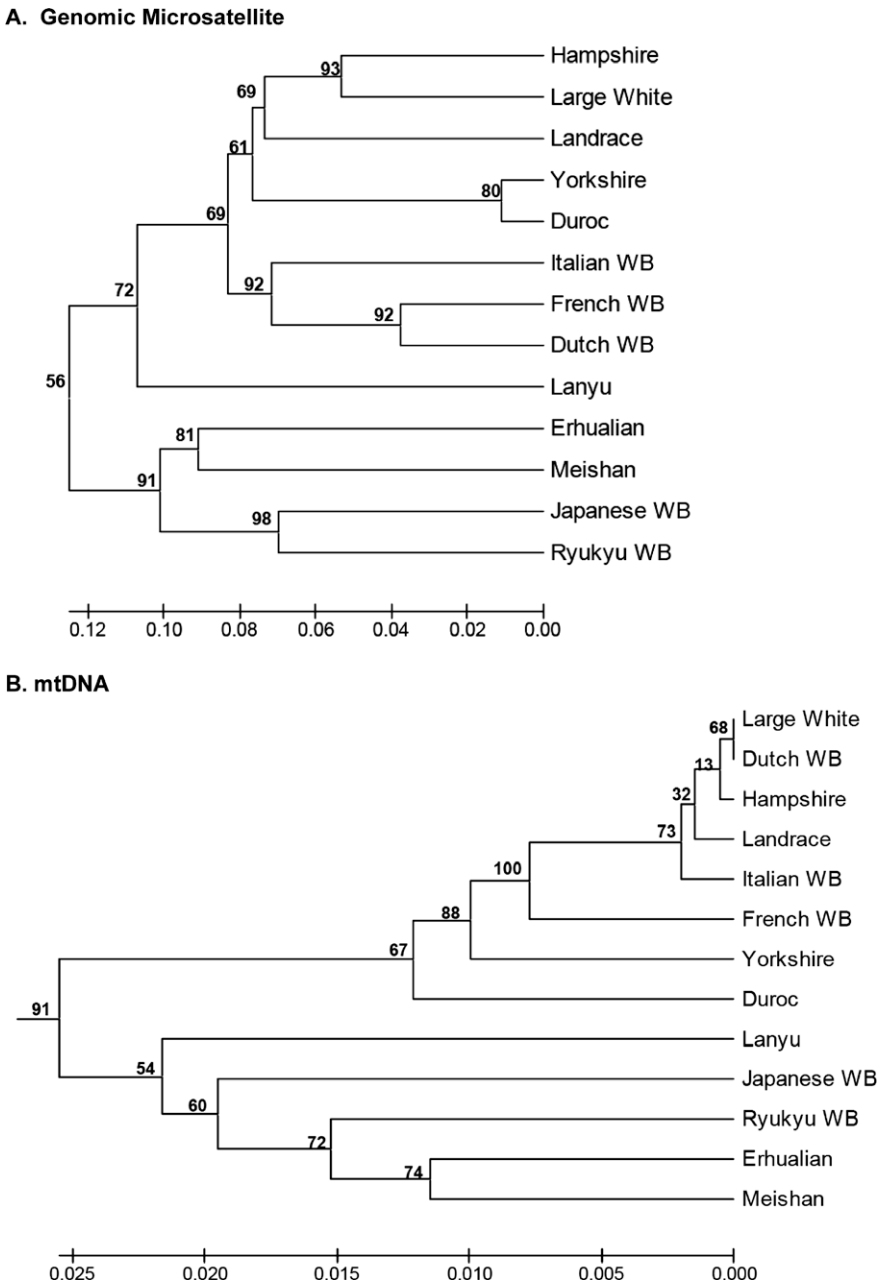

Fig. 2. UPGMA tree of genetic relatedness. UPGMA trees showing the genetic relationships among populations are based on Nei's genetic distances (A) using microsatellite markers (B) using D-loop mtDNA. The number in the branches indicates the percentage occurrence in 1000 bootstrap replicates and the scale indicates genetic distance.

the European domestic population. The tight clustering pattern was not observed within the Asian domestic population (Fig. 3). Although the first two principal components comprised only 44% of the total variation observed, they still revealed genetic differentiation among populations, especially within geographic regions.

4. Discussion

4.1. Intra-population genetic variation

The mean expected heterozygosity of the 13 sample groups ranged from 0.53 (Hampshire) to 0.76 (French wild boar) (Table 1). The Hampshire breed had the lowest level of expected heterozygosity followed by Large White, Meishan, Erhualian, Lanyu, Duroc, Italian wild boar, Landrace, Dutch wild boar, Yorkshire, Japanese and Ryukyu wild boar and the French wild boar sample groups. The expected levels of heterozygosity were generally higher than those reported previously for Asian and European populations by Fan et al. (2002) and Laval et al. (2000). The high degree of genetic variability identified in the present study most likely reflects the inclusion of diverse unrelated animals and the use of highly polymorphic MS designed for genetic management of groups. The number of effective alleles ranged from 2.29 (Hampshire) to 3.74 (French wild boar). The highest number of effective alleles was identified in the French and Ryukyu wild boar sample group, closely followed by the Dutch wild boar. The number of effective alleles in the domestic populations most likely reflects the result of a loss of genetic variation due to a limited population size of captive populations (Fang and Andersson, 2006). The order of sample groups from the lowest level to highest level of diversity was consistent when the expected heterozygosity and the number of

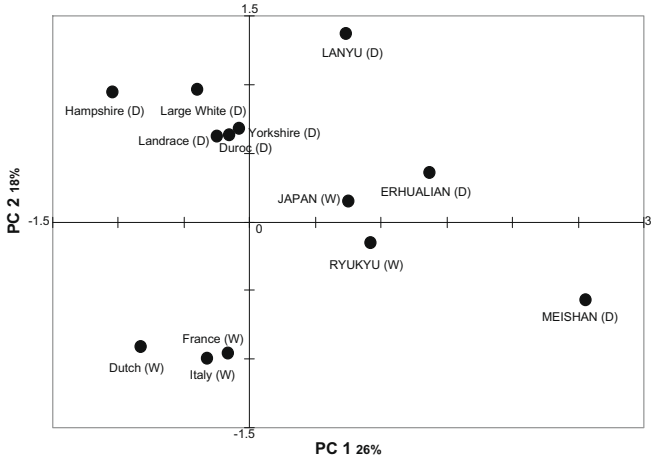

Fig. 3. Principal component analysis of pig populations. Scatter diagram showing relative positions of pig populations defined by principal component factor scores based on correlation matrix from allele frequency of microsatellite loci. The first principal component accounted for 26% and the second principal component accounted for 18% of the total variation. Asian pig populations are represented in all capital letters. (W) wild boar subpopulations; (D) domestic pig breeds.

effective alleles were compared. The French and Ryukyu wild boars consistently had the highest level of diversity, whereas the Italian wild boars had lower levels of diversity than the Landrace and Yorkshire breeds. The degree of genetic differentiation between European wild boars from those of Sardinia, most likely reflects the geographic barrier of the European Alps during glaciations (Kijas and Andersson, 2001). Prior to the seventh century C.E., pigs were not known to be located on the Mediterranean islands of Corsica and Sardinia. Thus, "wild boars" found on the island of Sardinia are most likely the result of human-aided introduction of Italian wild boars to the island (Larson et al., 2005). Thus, in the present study, lower levels of observed diversity in the Italian wild boars as compared to other Eurasian wild boar populations may be the result of geographic separation.

4.2. Inter-population genetic variation

Consistent with previous studies performed using mtDNA (Fang and Andersson, 2006), a genetic distinction was identified between European and Asian populations. Genetic separation was demonstrated by the overall F_{ST} estimate of 0.152 for Asian populations compared to European populations (0.250). The F_{ST} estimates identified were comparable to F_{ST} values of 0.18 found within Chinese indigenous swine populations and 0.27 in European domestic populations (Fan et al., 2002; Laval et al., 2000; Martinez et al., 2000). The purpose of this analysis was to further assess the structure of genetic variation in global pig populations. F_{ST} values up to 0.05 indicate negligible genetic differentiation whereas $\geqslant$0.25 provides evidence of genetic differentiation within the analyzed population (Kim et al., 2005). Thus, the estimations of genetic differentiation from the current study suggest that significant population subdivision exists over European and Asian pig populations.

Consistent with previous findings, the degree of genetic differentiation between populations within geographic regions was greatest between the Asian domestics and wild boar populations (F_{ST}: 0.29) as compared to the European domestics and wild boar populations (F_{ST}: 0.194) (Fan et al., 2002; Laval et al., 2000; Martinez et al., 2000). Using mtDNA, a previous study had revealed that despite geographic separation, selected pig populations in China, Korea and Japan were genetically more similar to each other than to populations from within the same country (Kim et al., 2002). In contrast, the current study revealed that the largest genetic differentiation was observed within the Asian domestic population suggesting that Asian domestic breeds were derived from distinct ancestral germplasms with little introgression between populations, which is in agreement with Megens et al. (2008). In comparison to the observations of Larson et al. (2005), the European domestic population in the current study revealed F_{ST} values less than 0.19 suggesting that within Europe many of the populations have experienced either mixing of germplasms or a shared single recent ancestor (Fig. 1). In contrast, the Asian domestic pig population may be derived from multiple sources in Asia. In the current study it is not possible to separate the hypotheses of mixing of germplasms and a single recent ancestor as the relatedness of past generations were not known.

4.3. Inter-population structures

Both the UPGMA and NJ phylogenetic trees using Nei's genetic distances, the bayesian analyses, maximum likelihood and the PCA showed distinct separation of the European and the Asian populations. The MS analyses revealed that the European domestic population clustered closely, while the wild boars clustered separately (Fig. 2A and Suppl. Fig. IA). The mitochondrial analysis provided the same pattern of separation between European and Asian populations with the exception of the Lanyu breed.

Consistent with previous studies utilizing mtDNA analyses (Larson et al., 2005; Huang et al., 1999), both MS and mtDNA analyses in the current study, revealed a distinct separation of the Asian domestic breeds (Meishan and Erhualian) from the Asian wild boar subpopulations. The Meishan and Erhualian breeds closely clustered together, which is in accordance with the developmental history of these Asian breeds. Both breeds originated in China and are classified as a subtype of the Taihu pig (Yang et al., 2003). Despite their genetic closeness, there is evidence that the Erhualian and Meishan may have originated from two different ancestral Taihu pigs; thus, resulting in two separate breeds (F_{ST}: 0.176) (Fan et al., 2002) and distinct phylogenetic branches, as is seen in the current study.

Conversely, from the MS analysis of this study based on Nei's genetic distances, the Lanyu population clustered more closely with the European populations than the Asian populations. The Lanyu pig is an indigenous pig breed inhabiting the Lanyu Islet, off the coast of Taiwan (Cheng, 1986). Morphologically the Lanyu pig differs from other Chinese breeds, but is thought to be closely related genetically. Wu et al. (2007) identified two distinct mtDNA haplotypes within Lanyu pig populations: Type I was independent of European and Asian pigs all together, while Type II clustered within the traditional Asian clade. From the mtDNA analysis done in the current study, the Lanyu breed clustered with the other Asian populations (Fig. 2B), consistent with Type II haplotypes. Alternatively, the MS phylogenetic tree in the current study placed the Lanyu breed as a separate distinct branch off the European branches, representing the presence of some European ancestral origins (Fig. 2A). The STRUCTURE cluster analysis results further supported the influence of both Asian and European origins in the development of the Lanyu breed with some individuals clustering over 50% with the European domestic population while others only contained Asian wild boar genomic haplotypes (Fig. 1). The presence of both Asian and European origins in the Lanyu animals, as seen in the mtDNA analysis, may suggest that the Lanyu breed was derived from Asian maternal origin, but had European genomic influences. The PCA plot implies that the Lanyu breed is genetically distinct from the Asian populations since it does not cluster closely with other Asian domestic or wild boar populations and may reflect a separate domestication event (Megens et al., 2008) (Fig. 3). The genetic separation of the Lanyu breed from other Asian domestic breeds could additionally be explained by the introgres-

sion of multiple boars with a single maternal origin. This observation is in contrast to a few studies done solely using mtDNA analyses, where they identified a single mtDNA haplotype for all Chinese/Asian domestic pigs (Huang et al., 1999).

5. Conclusions

Genomic MS markers provided a powerful tool to differentiate pig populations. This differentiation was observed with the clear separation not only of European and Asian populations (Fig. 2A) but also further detailed separation within each individual region. Although, in contrast to previous mtDNA analyses (Fang and Andersson, 2006), the STRUCTURE cluster analysis of the current study clearly revealed little introgression of Asian populations into European domestic breeds, if any. Together mtDNA and MS analyses provided further evidence as to the maternal and paternal origins of European and Asian domestic populations. The European domestic breeds maintained low pairwise F_{ST} values, suggesting little population differentiation between them. Additionally, close clustering identified in the UPGMA phylogenetic tree, STRUCTURE analysis and PCA plot suggests that the European domestic breeds were derived from a single ancestral origin. Conversely, larger pairwise F_{ST} values and distinct independent clustering identified in the UPGMA phylogenetic tree, STRUCTURE analysis and PCA plot suggest that there are multiple origins for Asian domestic breeds. The current findings support the hypothesis that Asian domestic populations were derived from multiple Asian ancestral origins whereas the European domestic populations represent a single ancestral European lineage.

Acknowledgments

We thank Swine Genetics International, Lieske Genetics, Birchwood Genetics Inc., United Hog Systems Inc., High Point Swine Genetics Inc., Infigen, Taiwan Livestock Research Institute, Beijing Agriculture University (BAU), and United Swine Genetics for providing domestic samples. Additionally we thank Dr. S. Casu of Istituto Zootecnico per la sardegna, Dr. G. Semiadi of Puslit Biologi, Dr. Okumura of Society for Techno-innovation of Agriculture, Forestry and Fisheries, Dr. A. Ducos of UMR INRA-ENVT de cytogénétique des populations animals, Dr. Groenen and Dr. Crooijmans of Wageningen University and Dr. H.P. Koelewijn of WUR for providing wild boar DNA. This work was supported in part by USDA/NRI-CSREES Grant AG2001-35205-11698, USDA-ARS and AG58-5438-2-313. E.L. was supported by the Initiative for Future Agriculture and Food Systems Grant No. 2001-52100-11527 from the USDA Cooperative State Research, Education, and Extension Service.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.11.004.

References

Berggren, K.T., Ellegren, H., Hewitt, G.M., Seddon, J.M., 2005. Understanding the phylogeographic patterns of European hedgehogs, *Erinaceus concolor* and *E. europaeus* using the MHC. Heredity 95, 84–90.

Bruford, M.W., Bradley, D.G., Luikart, G., 2003. DNA markers reveal the complexity of livestock domestication. Nature Reviews Genetics 4, 900–910.

Cheng, P., 1986. Pig breeds in China. In: Zhang, Z. (Ed.), Description of Chinese Pig Breeds. Shanghai Scientific and Technical Publishers, Shanghai, pp. 150–155.

Clutton-Brock, J., 1987. Man-made animals: pigs. In: A Natural History of Domesticated Mammals. Cambridge University Press, Cambridge, UK, pp. 70–77.

Corander, J., Waldmann, P., Marttinen, P., Sillanpaa, M.J., 2004. BAPS 2: enhanced possibilities for the analysis of genetic population structure. Bioinformatics 20 (15), 2363–2369.

Estoup, A., Solignac, M., Cornuet, J.M., Goudet, J., Scholl, A., 1996. Genetic differentiation of continental and island population of *Bombus terrestris* in Europe. Molecular Ecology 5, 19–31.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE a simulation study. Molecular Ecology 14, 2611–2620.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47.

Fan, B., Wang, Z., Li, Y., Zhao, X., Liu, B., Zhao, S., et al., 2002. Genetic variation analysis within and among Chinese indigenous swine populations using microsatellite markers. Animal Genetics 33, 422–427.

Fang, M., Andersson, L., 2006. Mitochondrial diversity in European and Chinese pigs is consistent with population expansions that occurred prior to domestication. Proceedings of the Royal Society B: Biological Sciences 273, 1803–1810.

Forbes, S.H., Hogg, J.T., Buchanan, F.C., Crawford, A.M., Allendorf, F.W., 1995. Microsatellite evolution congeneric mammals: domestic and bighorn sheep. Molecular Biology and Evolution 12, 1106–1113.

Giuffra, E., Kijas, J.M.K., Amarger, V., Carlborg, O., Jeon, J.-T., Andersson, L., 2000. The origin of the domestic pig: independent domestication and subsequent introgression. Genetics 154, 1785–1791.

Hammond, K., Leitch, H.W., 1998. Genetic resources and the global program for their management. In: Rothschild, M.F., Ruvinsky, A. (Eds.), The Genetics of the Pig. CAB International, New York, pp. 405–425.

Huang, Y.-F., Shi, X.-W., Zhang, Y.-P., 1999. Mitochondrial genetic variation in Chinese pigs and wild boars. Biochemical Genetics 37, 335–343.

Jonsson, P., 1991. Civilization and domestication of pigs. In: Maijala, K. (Ed.), Genetic Resources of Pig, Sheep and Goats. Elsevier, pp. 11–50.

Kijas, J.M.H., Andersson, L., 2001. A phylogenetic study of the origin of the domestic pig estimated from the near-complete mtDNA genome. Journal of Molecular Evolution 52, 302–308.

Kim, K.-I., Lee, J.-H., Li, K., Zhang, Y.-P., Lee, S.-S., Gongora, J., Moran, C., 2002. Phylogenetic relationship of Asian and European pig breeds determined by mitochondrial DNA D-loop sequence polymorphism. Animal Genetics 33, 19–25.

Kim, T.H., Kim, K.S., Choi, B.H., Yoon, D.H., Jang, G.W., Lee, K.T., Chung, H.Y., Lee, H.Y., Park, H.S., Lee, J.W., 2005. Genetic structure of pig breeds from Korea and China using microsatellite loci analysis. Journal of Animal Science 83, 2255–2263.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: Molecular Evolutionary Genetics Análisis software. Bioinformatics 17, 1244–1245.

Kurosawa, Y., Tanaka, K., Suzuki, S., Oishi, T., 1984. Variations of blood groups observed in wild pig populations in Japan. Japanese Journal of Zootechnical Science 55, 209–212.

Larson, G., Dobney, K., Albarella, U., Fang, M., Matisoo-Smith, E., Robins, J., et al., 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science 307, 1618.

Laval, G., Iannuccelli, N., Legault, C., Milan, D., Groenen, M., Giuffra, E., et al., 2000. Genetic diversity of eleven European pig breeds. Genet. Sel. Evol. 32, 187–203.

Li, M.D., Enfield, F.D., 1989. A characterization of Chinese breeds of swine using cluster analysis. Journal of Animal Breeding and Genetics 106, 379–388.

Martinez, A.M., Delgado, J.V., Rodero, A., Vega-Pla, J.L., 2000. Genetic structure of the Iberian pig breed using microsatellites. Animal Genetics 31, 295–301.

Megens, H.J., Crooijmans, R.P., San Cristobal, M., Hui, X., Li, N., Groenen, M.A., 2008. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. Genetics Selection Evolution 40 (1), 103–128.

Oishi, T., Amano, T., Tanaka, K., 1993. Phylogenetic relationship among seventeen pig breeds analyzed by principal component analysis based on blood groups and biochemical polymorphisms. Journal of Animal Genetics 21, 29–31.

Okumura, N., Kurosawa, Y., Kobayashi, E., Watanobe, T., Ishiguro, N., Yasue, H., et al., 2001. Genetic relationship amongst the major non-coding regions of mitochondrial DNAs in wild boars and several breeds of domesticated pigs. Animal Genetics 32, 139–147.

Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Molecular Biology and Evolution 5 (5), 568–583.

Pritchard, J.K., Stephens, P., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Raymond, M., Rousset, F., 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86, 248–249.

Weir, B.S., Cockerham, C.C., 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38, 1358–1370.

Wu, C.Y., Jiand, Y.N., Chu, H.P., Li, S.H., Wang, Y., Li, Y.H., et al., 2007. The type I Lanyu pig has a maternal genetic lineage distinct from Asian and European pigs. Genetics Selection Evolution 38, 499–505.

Yang, S., Wang, Z., Liu, B., Zhang, G., Zhao, S., Yu, M., et al., 2003. Genetic variation and relationships of eighteen Chinese indigenous pig breeds. Genetics Selection Evolution 35, 657–671.

Yeh, F.C., Boyle, T., Rongcai, Y., Ye, Z., Xin, J.M., 1999. POPGENE, Version 1.31. A Microsoft Window Based Free Ware for Population Genetic Analysis. University of Alberta, Edmonton, AB.

Yue, G.H., Beeckmann, P., Barternschlager, H., Moser, G., Geldermann, H., 1999. Rapid and precise genotyping of porcine microsatellites. Electrophoresis 20 (17), 3358–3363.

Zhang, Z.G., 1986. The Pig Breeds in China. Shanghai Scientific and Technical Publishers, Shanghai.