Integrating Y-Chromosome, Mitochondrial, and Autosomal Data to Analyze the Origin of Pig Breeds

O. Ramírez,[1] *A. Ojeda,** *A. Tomàs,** *D. Gallardo,** *L.S. Huang,†* *J.M. Folch,** *A. Clop,‡* *A. Sánchez,** *B. Badaoui,** *O. Hanotte,§* *O. Galman-Omitogun,‖* *S.M. Makuza,¶* *H. Soto,#* *J. Cadillo,**** *L. Kelly,†††* *I.C. Cho,‡‡* *S. Yeghoyan,§§* *M. Pérez-Enciso,*‖‖* *and M. Amills**

*Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, Bellaterra, Spain; †Key Laboratory for Animal Biotechnology of Jiangxi Province and the Ministry of Agriculture of China, Jiangxi Agricultural University, Nanchang, China; ‡Molecular and Medical Genetics Department, King's College London, London, United Kingdom; §International Livestock Research Institute, P.O. Box 30709, Nairobi, Kenya; ‖Department of Animal Science, Faculty of Agriculture, Obafemi Awolowo University, Ile-Ife-Nia, Nigeria; ¶Department of Animal Science, University of Zimbabwe, Harare, Zimbabwe; #Escuela de Zootecnia, Facultad de Ciencias Agroalimentarias, Universidad de Costa Rica, San Pedro, Costa Rica; **Departamento de Producción Animal, Facultad de Zootecnia, Universidad Nacional Agraria La Molina, La Molina, Peru; ††Estación Experimental, INIA Las Brujas, Canelones, Instituto Nacional de Investigaciones Agrarias, Uruguay; ‡‡Animal Genomics and Bioinformatics Division, National Institute of Animal Science, RDA, Kwonsun-Gu, Suwon, South Korea; §§Department of Epidemiology and Parasitology, Armenian State Agrarian University, Yerevan, Armenia; and ‖‖Institut de Recerca i Estudis Avançats de Catalunya, ICREA, Barcelona, Spain

We have investigated the origin of swine breeds through the joint analysis of mitochondrial, microsatellite, and Y-chromosome polymorphisms in a sample of pigs and wild boars with a worldwide distribution. Genetic differentiation between pigs and wild boars was remarkably weak, likely as a consequence of a sustained gene flow between both populations. The analysis of nuclear markers evidenced the existence of a close genetic relationship between Near Eastern and European wild boars making it difficult to infer their relative contributions to the gene pool of modern European breeds. Moreover, we have shown that European and Far Eastern pig populations have contributed maternal and paternal lineages to the foundation of African and South American breeds. Although West African pigs from Nigeria and Benin exclusively harbored European alleles, Far Eastern and European genetic signatures of similar intensity were detected in swine breeds from Eastern Africa. This region seems to have been a major point of entry of livestock species in the African continent as a result of the Indian Ocean trade. Finally, South American creole breeds had essentially a European ancestry although Asian Y-chromosome and mitochondrial haplotypes were found in a few Nicaraguan pigs. The existence of Spanish and Portuguese commercial routes linking Asia with America might have favored the introduction of Far Eastern breeds into this continent.

Introduction

The last decade has witnessed considerable advances in our understanding of the historical events that led to pig domestication and breed formation. Since the pioneer study of Giuffra et al. (2000), establishing the existence of two main centers of domestication, a wide array of Asian and European breeds has been analyzed with mitochondrial markers (Larson et al. 2005; Fang and Andersson 2006; Wu et al. 2007) and microsatellites (Fang et al. 2005; SanCristobal et al. 2006; Megens et al. 2008). This research has evidenced a deep population split between Asian and European populations, with an estimated time of divergence that ranges from 58,000 (Kim et al. 2002) to 900,000 YBP (Giuffra et al. 2000; Larson et al. 2005; Fang and Andersson 2006). Moreover, large-scale mitochondrial analysis of pigs and wild boars with a worldwide distribution revealed that pigs were domesticated at multiple locations across Eurasia (Larson et al. 2005; Wu et al. 2007).

One limitation of previous studies lies in the fact that they are exclusively based on one type of marker (either mitochondrial DNA or microsatellites), instead of integrating multiple sources of genomic information. Moreover, the analysis of the patrilineal history of pig breeds has not been undertaken so far due to the lack of Y-chromosome markers. The general goal of the current work was to investigate the origins of pig breeds through the simultaneous analysis of three different sources of genomic data (i.e., mitochondrial, Y-chromosome, and autosomal polymorphisms) in wild and domesticated *Sus scrofa* populations with a worldwide distribution. In this framework, the comparison of genetic diversity between pigs and wild boars would be particularly meaningful in order to understand how domestication and artificial selection have shaped the allelic repertoire of swine breeds. Moreover, we aimed to address two specific questions of broad interest. The absence of Near Eastern mitochondrial haplotypes in modern European breeds has been taken as evidence that they descend from pigs domesticated locally (Larson et al. 2005). We wanted to find out if autosomal markers, which have a much lower extinction rate than their mitochondrial counterparts (Zhang and Hewitt 2003), support this notion or not. In addition, we were interested in investigating the ancestry of African and South American pig breeds, two populations that have not been previously analyzed. Africa and South America were explored and colonized by the Europeans in the 15th century onward (Ferro 1997). This circumstance might lead us to assume that African and South American pigs have fundamentally a European origin. However, the participation of Asian breeds cannot be ruled out because the ancient introduction of Indian and Far Eastern livestock in the Eastern coast of Africa has been widely documented (Hanotte et al. 2002; Muchadeyi et al. 2008). Moreover, several Brazilian pig breeds, such as Nilo and Canastrinho, show a close phenotypic resemblance with Asian pigs (Porter 1993). These facts brought us to

[1] Present address: Institute of Evolutionary Biology (UPF-CSIC), Barcelona, Spain.

Key words: domestication, pig, wild boar, phylogeography, Y-chromosome, mitochondrial DNA, microsatellite, *Sus scrofa*, admixture.

E-mail: Marcel.Amills@uab.es.

Mol. Biol. Evol. 26(9):2061–2072. 2009
doi:10.1093/molbev/msp118
Advance Access publication June 17, 2009

investigate if the patterns of genetic variation observed in African and South American pig breeds are more consistent either with a single European origin or, conversely, with a more complex model integrating European and Asian influences.

Materials and Methods

Animal Material

Animal material employed for the analysis of the autosomal genome comprised 374 pigs and wild boars with a wide geographical distribution including Europe (203), Near East (22), Africa (71), America (30), and Asia (48), as shown in supplementary table 1, Supplementary Material online. A subset of 266 pigs and wild boars (supplementary table 1, Supplementary Material online) was used for analyzing the polymorphism of *Sus scrofa* Y-chromosome (SSCY). The data set employed for mitochondrial cytochrome b (*MT-CYB*) analysis included 345 sequences (190 characterized in the current work and 155 extracted from the GenBank) and it is described in supplementary table 2, Supplementary Material online. *Sus barbatus* and *Babyrousa babyrussa* were employed as outgroups in SSCY, *MT-CYB*, and microsatellite analyses.

Pig genomic DNA was isolated from either blood or tissues by using a conventional phenol–chloroform precipitation protocol, as described elsewhere. In contrast, nucleic acid isolation from hair shafts was achieved by using the DNeasy Blood and Tissue Kit (Qiagen, Barcelona, Spain). The only procedure we have modified in this commercial protocol was the lysis method, which involved an incubation in 350 μl lysis buffer (100 mM Tris–Cl pH = 8, 100 mM NaCl, 3 mM Ca Cl_2, 2% sodium dodecyl sulfate, 40 mM dithiotreitol, and 250 μg/ml proteinase K) at 56 °C for 3–4 h. Samples that yielded a poor amount of DNA were amplified at the whole-genome level with the GenomiPhi DNA Amplification Kit (Amersham Biosciences Europe GmbH, Cerdanyola del Vallès, Spain) according to the manufacturer's instructions.

Y-Chromosome Sequencing and Genotyping

In order to identify polymorphic sites in the SSCY, we have sequenced 2,328 bp of SSCY in 28–39 wild boars and pigs from Europe (11–18), Near East (2), Africa (4–8), and Asia (8–12). Sequences of primers used for amplifying and sequencing each one of the SSCY-targeted regions are reported in supplementary table 3, Supplementary Material online. More specifically, we have sequenced 427-bp intron 24 of the ubiquitin-specific protease 9 (*USP9Y*) gene (GenBank accession number: EU549792–94), one 543-bp region located in the amelogenin (*AMELY*) gene (GenBank accession number: EU549795–99), 184 bp of eukaryotic translation initiation factor 2, subunit 3 (*EIF2s3Y*) locus (GenBank accession number: AJ437581), and three fragments corresponding to the ubiquitously transcribed tetratricopeptide repeat gene (*UTY*) gene. These three amplicons corresponded to 330-bp intron 1 (GenBank accession number: EU549788–91), 491-bp intron 7 (GenBank accession number: EU549785–87), and 353-bp intron 9 (GenBank accession number: EU549800–01) of the *UTY* gene.

These six polymerase chain reactions (PCRs) were carried out in a 25-μl final reaction mixture containing PCR buffer (1×), 2.5 mM $MgCl_2$, 0.2 mM dNTPs (Applied Biosystems, Warrington, United Kingdom), 0.5 μM of each primer, and 0.75 U *TaqGold* DNA polymerase (Applied Biosystems). Thermocycling profiles were one cycle at 95 °C for 10 min, followed by 35 cycles of 94 °C for 1 min, T_m (see supplementary table 3, Supplementary Material online) for 1 min and 72 °C for 1 min and a final extension step at 72 °C for 15 min. Amplified products were purified with the ExoSAP-IT kit [Amersham Biosciences Europe GmbH, Cerdanyola del Vallès, Spain]) and polymorphisms were genotyped by primer-extension analysis by using the SnaPshot ddNTP Primer-Extension kit (Applied Biosystems). With the aim of checking the specificity of the PCRs (lack of cross-amplification of X-linked or autosomal loci), six independent control PCRs were performed by employing female genomic DNA as a template. The lack of an amplified product was a necessary requirement to validate each PCR.

Microsatellite Genotyping

A total of 12 autosomal and unlinked microsatellites were analyzed: S0155 (chromosome 1), SW240 (chromosome 2), S0090 (chromosome 2), SW72 (chromosome 3), SW911 (chromosome 9), SW951 (chromosome 10), S0386 (chromosome 11), SW857 (chromosome 14), SW936 (chromosome 15), S0355 (chromosome 15), S0101 (chromosome 17), and SW24 (chromosome 17). Microsatellites were chosen based on their ease of scoring, absence of null alleles, location and informativeness. Microsatellites were amplified using three multiplex PCRs : multiplex-1 (SW911, SW857, SW240, and S0090), multiplex-2 (SW936, SW72, and S0155), and multiplex-3 (S0101, S0355, S0386, SW951, and SW24). PCRs were carried out in a 10-μl reaction mixture containing PCR buffer (1×), 1.5 mM $MgCl_2$, 0.25 mM of each deoxynucleotide, 0.5 U *TaqGold* DNA polymerase (Applied Biosystems), and 30–40 ng genomic DNA. Primer concentrations were optimized for each marker: 0.35 μM for S0155, 0.3 μM for S0090 and SW24, 0.25 μM for SW911 and SW72, 0.15 μM for SW240 and SW936, and 0.2 μM for the remaining microsatellites. Thermocycling profiles were 10 min at 94 °C followed by 35 cycles of 94 °C (30 s), 58 °C (1 min), and 72 °C (1 min), followed by a final extension step at 72 °C for 7 min. Amplified products were electrophoresed in an ABI Prism 3730 Genetic Analyzer equipment with fluorescent detection (Applied Biosystems) and analyzed with the GeneScan 3.7 software (Applied Biosystems).

Mitochondrial Sequencing

We have analyzed a 895-bp fragment corresponding to the mitochondrial *MT-CYB* gene. Primer sequences and amplification profiles have been described by Alves et al. (2003). PCR products were purified with the ExoSAP-IT kit (Amersham Biosciences Europe GmbH, Cerdanyola del Vallès, Spain) and sequenced with the Big Dye Terminator Cycle Sequencing Ready Reaction kit v3.1 (Applied

Table 1
Description of *Sus scrofa* Y-Chromosome Haplotypes (HY1, HY2, and HY3)

	UTY				*USP9Y*	*AMELY*	
	Intron 1			Intron 9	Intron 24	5′ End	
Haplotype	G/C	T/C	C/G	C/G	T/C	A/G	C/T
HY1	G	T	C	C	T	A	C
HY2	G	T	C	C	T	G	C
HY3	C	C	G	G	C	G	T
HY4[a]	G	C	C	G	C	G	T

[a] *Sus barbatus* and *Babyrousa babyrussa* were used as outgroups (HY4).

Biosystems). Sequencing reactions were analyzed by capillary electrophoresis in an automated ABI PRISM 3730 capillary electrophoresis device (Applied Biosystems). Sequences were aligned using the SeqScape software v2.6 (Applied Biosystems).

Data Analysis

Nucleotide diversities (π) were computed with the DnaSP 4.90.1 software (Rozas et al. 2003). The median-joining network was constructed with the Network 4.5 program (Bandelt et al. 2000), whereas the Bayesian phylogenetic tree was built with MrBayes v3.1.2 (Ronquist and Huelsenbeck 2003). In this bayesian analysis, we employed a Hasegawa, Kishino, and Yano evolutionary model (Hasegawa et al. 1985) previously selected with Modeltest (Posada and Crandall 1998). Evolutionary parameters (proportion of invariant sites, α-value and transition to transversion rates) were estimated with both Modeltest and MrBayes. The consensus tree was visualized with TreeView (Page 1996) and edited with TreeDyn (Chevenet et al. 2006).

The Time to Most Recent Common Ancestor (TMRCA) of SSCY haplotypes was estimated with the Genetree program (http://www.stats.ox.ac.uk/~griff/software.html). Genetree provides mutation-age estimates as multiples of θ; thus, either effective population size (N_e) or mutation rate (μ) should be fixed a priori to transform time expressed in θ units to a given number of generations. In the absence of a pig-specific nucleotide substitution rate for the Y-chromosome, we employed the one reported for human that is approximately 1.65×10^{-9} per bp and year (Schaffner 2004). This value has been shown to be fairly constant among mammals (Kumar and Subramanian 2002). All Genetree program executions were run for 1,000,000 iterations.

Expected and observed heterozygosities within each population were calculated for each microsatellite using GENEPOP 3.1 (Rousset and Raymond 1995). In addition, all genotypes were screened using a Bayesian admixture procedure implemented in the STRUCTURE 2.2 software (Pritchard et al. 2000). STRUCTURE was run with 10^6 iterations, following a burn-in period of 10,000 iterations, to estimate the number of populations (K) using only genetic information. To infer the correct K value, we applied the ΔK method (Evanno et al. 2005). We estimated the levels of European versus Far Eastern genetic admixture in five pig populations (South American, African, Mediterranean, International, and Anglo-Saxon swine) by analyzing autosomal microsatellite data with the Leadmix software. We assumed that European wild boars and Far Eastern wild boars and pigs are the two ancestral populations. Leadmix analysis was performed in accordance with the recommendations of Wang (2003). Robert and Hiorns (RH; 1965), Long and Chakraborty (LC; 1991, 1992), and Wang (W; 2003) estimators were used to assess the levels of population admixture. The 95% confidence intervals for RH and LC moment estimators were obtained by bootstrapping (1,000 replicates over loci), whereas those corresponding to the W estimator were calculated through the analysis of profile log-likelihood curves.

Analysis of the molecular variance (AMOVA) was carried out by using the Arlequin software (Excoffier et al. 2005) with the goal of estimating the proportion of genetic variation within and among groups (pig vs. wild boars) and populations. Arlequin was also employed to estimate F_{ST} statistics and their statistical significances (P values were calculated after performing 10,000 permutations). A tree based on D_A distances (Nei et al. 1983) was constructed by using the Neighbor-Joining clustering algorithm (Saitou and Nei 1987). The distance matrix was obtained with the program POPULATIONS v1.2.28 (Langella 2002) and the tree was built with the MEGA 4.1 package (Tamura et al. 2007).

Results

Polymorphism of Pig Y-Chromosome

Partial sequencing of six Y-chromosome regions located in the *USP9Y*, *UTY*, *AMELY*, and *EIF2s3Y* single-copy genes was performed. Multiple alignment of sequences corresponding to European, Far Eastern, African, and Near Eastern wild boars and/or pigs (mean number of sequences per region: 34.75) revealed the existence of seven single nucleotide polymorphisms that grouped into three haplotypes HY1, HY2, and HY3 (table 1). The HY1 and HY2 haplotypes were closely related differing by a single substitution at the *AMELY* gene and both of them were highly divergent with regard to HY3 (mean difference: 2.7 substitutions/kb). Nucleotide diversity of SSCY was remarkably high in pigs ($\pi = 13.3 \times 10^{-4} \pm 2.9 \times 10^{-4}$) and wild boars ($\pi = 9.8 \times 10^{-4} \pm 5.2 \times 10^{-4}$) in spite of the low number of haplotypes. This high nucleotide diversity is mostly explained by the existence of two highly divergent and ancient lineages (HY1 and HY2 vs. HY3), with an estimated TMRCA of 1.27 My (0.49–2.05 My at 95% highest posterior density).

Geographical Distribution of Pig Y-Chromosome Haplotypes

Analysis of the geographical distribution of SSCY haplotypes in a panel of 266 male pigs and wild boars with a worldwide distribution revealed the existence of a well-defined geographical pattern in Eurasia (table 2, supplementary tables 4 and 5, Supplementary Material online). European and Near Eastern wild boars as well as European pigs displayed the HY1 haplotype at high frequencies (0.91–0.93), whereas HY3 was minority (0.00–0.03) being

Table 2
Frequencies of Y-Chromosome Haplotypes (HY1, HY2, and HY3) in 10 *Sus scrofa* Populations

Population	HY1	HY2	HY3	*N*
EWB	0.83	0.11	0.06	36
INTP	0.99	0.01	0.00	84
ANGLP	0.88	0.00	0.12	17
MEDLP	0.93	0.07	0.00	27
Overall Europe	0.93	0.04	0.03	164
AFWB	0.57	0.43	0.00	7
AFLP	0.42	0.00	0.58	31
Overall Africa	0.45	0.08	0.47	38
AWB	0.13	0.40	0.47	15
AP	0.68	0.00	0.32	19
Overall Asia	0.44	0.18	0.38	34
NEWB	0.91	0.09	0.00	11
SCAP	0.84	0.00	0.16	19
Total	0.79	0.06	0.14	266

N = number of genotyped individuals, EWB = European wild boar, INTP = International pigs, ANGLP = Anglo-Saxon local pigs, MEDLP = Mediterranean and Slav local pigs, AFWB = African wild boar, AFLP = African local pigs, NEWB = Near Eastern wild boar, SCAP = South and Central American local pigs, AWB = Far Eastern wild boar, and AP = Far Eastern local pigs.

exclusively found in two Tamworth pigs (this breed has been extensively introgressed with Asian alleles) and in two wild boars from Russia and the United Kingdom (supplementary tables 4 and 5, Supplementary Material online). Conversely, Far Eastern domestic pigs and wild boars showed similar frequencies for these two highly divergent lineages (HY1 + HY2: 0.53–0.68, HY3: 0.32–0.47). The same pattern was observed in African breeds (HY1 + HY2: 0.42, HY3: 0.58), with Kenyan and Zimbabwean (Mukota) pigs displaying HY3 at high frequencies (table 2, supplementary tables 4 and 5, Supplementary Material online). Finally, most of South American pigs carried exclusively the HY1 haplotype with the only exception of Nicaraguan and Argentinian pigs, where HY3 was detected (table 2, supplementary tables 4 and 5, Supplementary Material online). The differential distribution of the HY3 haplotype in Europe, where it is virtually absent, and East Asia reveals the absence of a male-mediated gene flow between these two centers of pig domestication from very ancient times.

Mitochondrial Diversity

Genetic analysis of *MT-CYB* variation (fig. 1, supplementary figs. 1 and 2, Supplementary Material online) confirmed the existence of two highly differentiated Far Eastern and European mitochondrial gene pools, as previously shown by other authors (Giuffra et al. 2000; Larson et al. 2005; Fang and Andersson 2006). Levels of genetic diversity were similar in pigs ($\pi = 88.3 \times 10^{-4} \pm 5.1 \times 10^{-4}$) and wild boars ($\pi = 70.6 \times 10^{-4} \pm 6.7 \times 10^{-4}$). We found 101 different mitochondrial *MT-CYB* haplotypes (Hap_1 to Hap_101) that are listed in supplementary table 2, Supplementary Material online. With the aim of establishing a correspondence with previous studies, we grouped these haplotypes in nine *MT-CYB* lineages named as E1, E2, E3, E4, A1, A2, A3, A4, and A5 (identified in the current work) following the nomenclature proposed by Fang and Andersson (2006). Although these authors consider E1 as a European lineage, according to our data it has a much wider geographical distribution including, among other territories, North Africa and Near East (supplementary table 2, Supplementary Material online). Consistent with data presented by Larson et al. (2005), Near Eastern *MT-CYB* sequences did not cluster with their European counterparts in the median-joining network (fig. 1) with the only exception of two sequences from Armenian wild boars. This finding suggests that European and Near Eastern gene pools have independent origins. With regard to

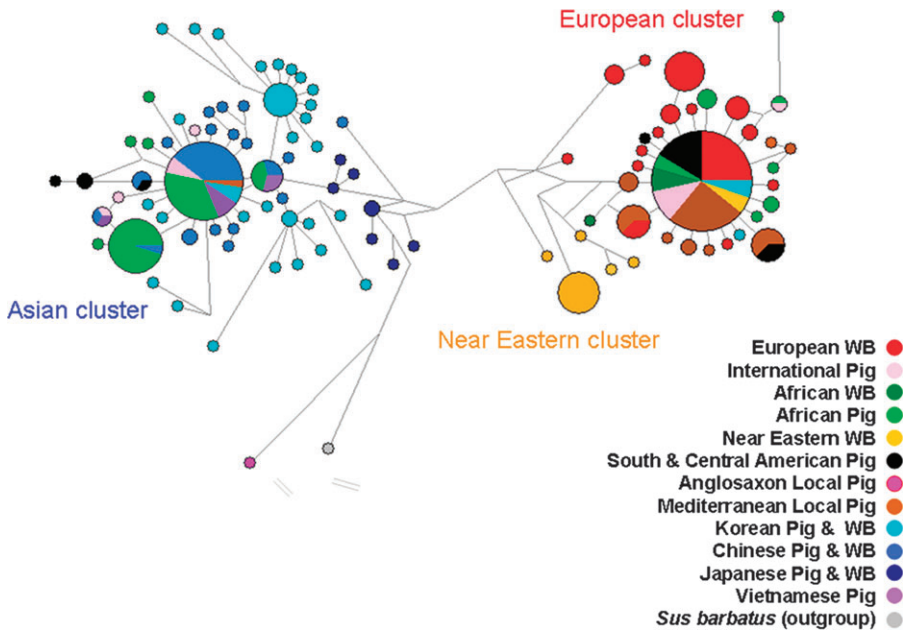

FIG. 1.—Median-joining network of 345 *MT-CYB* sequences corresponding to worldwide pig and wild boar (WB) populations. This analysis evidenced the existence of three divergent European, Near Eastern, and Asian clusters.

Table 3
Measurements of Expected (H_e) and Observed (H_o) Heterozygosities and Mean Number of Alleles (A) for a Panel of 12 Autosomal Microsatellites Genotyped in 374 *Sus scrofa* Individuals and two Outgroup Species (*Sus barbatus* and *Babyrousa babyrussa*)

	Populations[a]										
Genetic Diversity	EWB	AWB	AFWB	NEWB	SCAP	MEDLP	ANGLP	INTP	AFLP	AP	Outgroup[b]
H_e	0.638	0.814	0.651	0.707	0.734	0.719	0,734	0.703	0.748	0.752	0.79
H_o	0.536	0.55	0.58	0.602	0.576	0.507	0,478	0.574	0.641	0.564	0.583
A	7.833	8.5	3.833	6.917	8	7.5	6,333	7.333	9.5	7.583	2.833
N[c]	50.75	15	8.583	21.75	29.083	34.833	19	96.417	61.167	26	2

[a] EWB = European wild boar, INTP = International pigs, ANGLP = Anglo-Saxon local pigs, MEDLP = Mediterranean and Slav local pigs, AFWB = African wild boar, AFLP = African local pigs, NEWB = Near Eastern wild boar, SCAP = South and Central American local pigs, AWB = Far Eastern wild boar, and AP = Far Eastern local pigs.

[b] Outgroups: *Sus barbatus* and *Babyrousa babyrussa.*

[c] Mean number of analyzed individuals in each population.

the Far Eastern *Sus scrofa* specimens, the topology of the *MT-CYB* network happened to be remarkably complex. We identified a main cluster, that contained Chinese, Vietnamese Korean, and European pig *MT-CYB* sequences, and several others that were scattered throughout the network and contained sequences from Korean and Japanese wild boars (fig. 1, supplementary figs. 1 and 2, Supplementary Material online). The clustering of European *MT-CYB* sequences with the Far Eastern ones was expected given the extensive introgression of Chinese alleles into British breeds (Porter 1993), as previously shown by others (Giuffra et al. 2000; Larson et al. 2005; Fang and Andersson 2006).

In North (Tunisian and Moroccan wild boars) and West (pigs from Nigeria and Benin) Africa, Far Eastern haplotypes were completely absent, whereas they were particularly abundant in the Eastern part of the continent (Kenya and Zimbabwe, supplementary table 2, Supplementary Material online). With regard to South America, most pig samples displayed the H_1 haplotype (68%), whereas the second most frequent haplotype was H_11 (12%). These haplotype frequencies are very similar to the ones observed in the Iberian breed from Spain (H_1: 65%, H_11: 17%). In fact, the only European pigs that carried H_11 belonged to the Iberian breed (supplementary table 2, Supplementary Material online). These findings evidence the Iberian ancestry of South American pigs. Remarkably, Far Eastern *MT-CYB* haplotypes were found in Nicaraguan pigs. An additional median-joining network was constructed to delineate with more detail genetic relationships among Far Eastern, African, and American *MT-CYB* haplotypes (supplementary fig. 3, Supplementary Material online). This analysis showed that African sequences were distributed into four main clusters (from A to D), of which three contained Far Eastern *MT-CYB* sequences (A, B, and C), whereas the fourth (D) was clearly European and included, among others, African wild boar sequences (supplementary fig. 3, Supplementary Material online). In clusters A–C, African sequences grouped indistinctly with Indonesian, Chinese, Vietnamese, and Korean *MT-CYB* haplotypes making it difficult to discern their origin. Most of South American pig sequences grouped in cluster D, consistent with their European ancestry, whereas a number of them were closely related with the Far Eastern cluster A.

Variability of Autosomal Microsatellites

Analysis of 12 autosomal microsatellites showed that genetic diversity in each one of the analyzed *Sus scrofa* populations was fairly similar (table 3) and comparable with data presented in previous reports (Fang et al. 2005; SanCristobal et al. 2006; Megens et al. 2008). Genetic relationships among *Sus scrofa* populations were analyzed with the STRUCTURE software (Pritchard et al. 2000) as shown in figure 2. When the number of assumed populations (K) was set to three, we observed a red cluster containing wild boar populations from Europe, North Africa, and Near East as well as Mediterranean and South American pigs; a blue cluster with Far Eastern wild boars and pigs; and a green cluster including Anglo-Saxon, International, and African pig breeds. The existence of this green cluster is highly consistent with *MT-CYB* data indicating that African and International pig breeds have been formed by admixing the allelic pools of European and Far Eastern populations. In this way, this green cluster might be considered as an intermediate between the European (red) and Asian (blue) ones. This hypothesis was tested by using the Leadmix software (Wang 2003). We defined as parental populations those located in the two main centers of pig domestication, that is, European wild boars (population 1, $N = 52$) and a mixture of Far Eastern wild boars and pigs (population 2, $N = 48$). Previous microsatellite analyses had shown that these two populations are highly differentiated (Megens et al. 2008). Leadmix results are shown in table 4. The percentage of European alleles (averaged across methods) was high in Mediterranean (80%) and South American (67%) pigs, intermediate in Anglo-Saxon (53%) and African (49%) breeds, and relatively low in International swine (39%). Similar results were obtained when we considered as ancestral populations a mixture of European, Near Eastern, and African wild boars versus Far Eastern wild boars and pigs (data not shown). Finally, the STRUCTURE analysis indicated that European and Near Eastern wild boars are tightly related at the autosomal level because these two populations clustered together even when K was set to 8. This result did not match other analyses performed by us (fig. 1) and others (Larson et al. 2005) showing that European and Near Eastern wild boars are differentiated at the mitochondrial level.

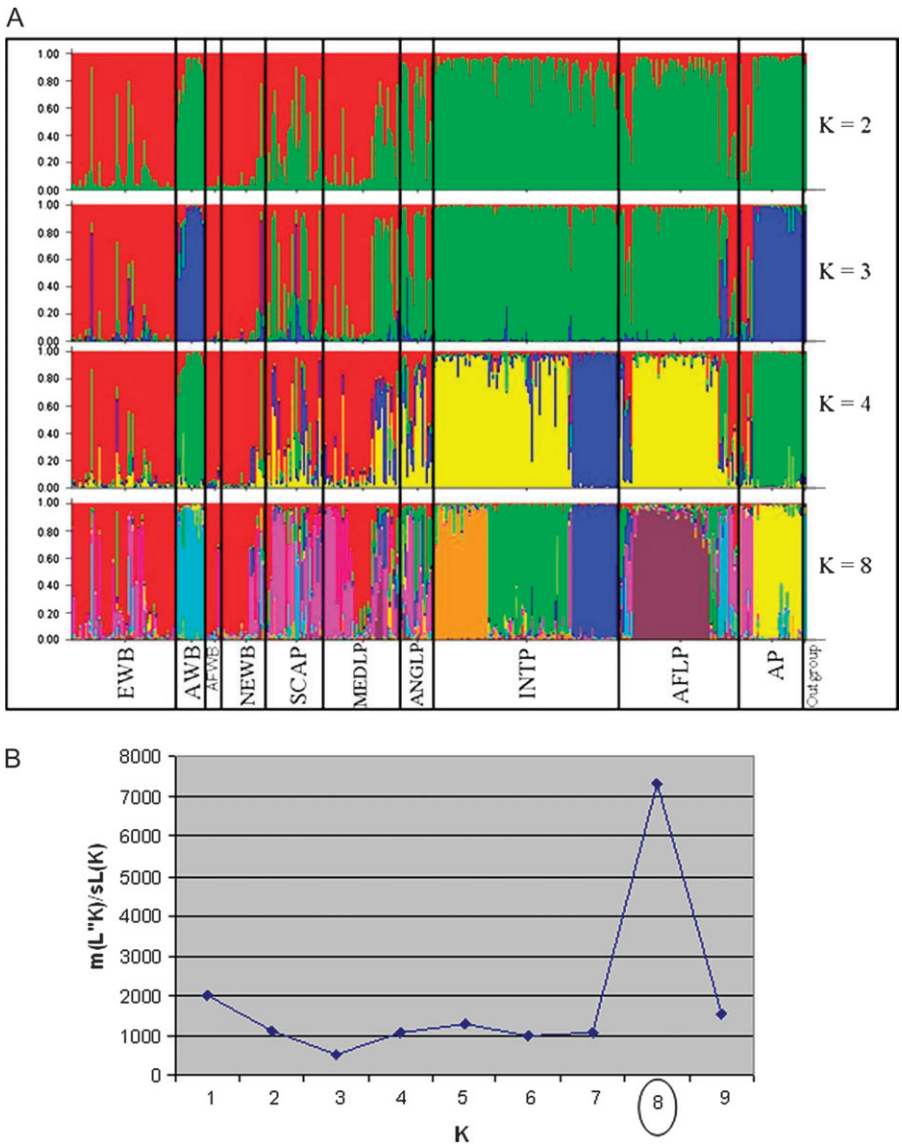

FIG. 2.—(*A*) Structure of 10 populations of wild boars and pigs ($N = 374$) and two outgroups. Each individual is represented by a single vertical line divided into 1–8 colors corresponding to a number K of clusters. The length of the colored segment in each vertical line shows the individual estimated fractional membership for each cluster. Black vertical lines separate the 10 analyzed *Sus scrofa* populations: EWB = European wild boar, AWB = Far Eastern wild boar, AFWB = African wild boar, NEWB = Near Eastern wild boar, SCAP = South and Central American local pigs, MEDLP = Mediterranean and Slav local pigs, ANGLP = Anglo-Saxon local pigs, INTP = International pig breeds, AFLP = African local pigs, and AP = Far Eastern local pigs. *Sus barbatus* and *Babyrousa babyrussa* were used as outgroups. Results shown are mean values obtained by averaging 20 STRUCTURE runs for each value of K. (*B*) Graphical method implemented by Evanno et al. (2005) in order to detect the true number of groups (K). The modal value of this distribution is the true K (encircled) or the uppermost level of structure (herewith, eight clusters).

As a complementary approach to investigate the genetic relationships among *Sus scrofa* populations, we calculated the corresponding F_{ST} values and their statistical significances. Although F_{ST} coefficients of genetic differentiation ranged from low ($F_{ST} = 0.03$) to high ($F_{ST} = 0.19$) values depending on the populations compared, all of them happened to be highly significant ($P < 1 \times 10^{-6}$). A multidimensional scaling plot of F_{ST} values and a Neighbor-Joining tree based on D_A distances are shown in figures 3A and 3*B*, respectively. Taken together, these analyses suggested that North African, European, and Near Eastern wild boars as well as Mediterranean pigs are clearly differentiated from Far Eastern pigs and wild boars, whereas Anglo-Saxon, International, and African pig breeds occupy an intermediate position between these two main groups. The multidimensional scaling plot and the Neighbor-Joining tree also showed the existence of a close relationship between Near Eastern and European wild boars (fig. 3), although it should be emphasized that the coefficient of genetic differentiation between these two populations although low ($F_{ST} = 0.08$) is highly significant ($P < 1.10^{-6}$)

Analysis of Population Structure

An AMOVA of mitochondrial, SSCY, and autosomal markers was performed to study *Sus scrofa* population structure

Table 4
Levels of European and Far Eastern Genetic Admixture in Five Pig Populations Estimated through the Analysis of Autosomal Microsatellites

Population[a]	Percentage of Admixture (% European Background)[b]		
	RH[c]	LC[c]	W[c]
SCAP	70.15 (54.40, 86.79)	58.45 (50.36, 72.37)	73.55 (62.15, 84.04)
MEDLP	74.95 (63.62, 85.57)	82,60 (73.39, 91.50)	81.22 (68.64, 92.04)
ANGLP	60.07 (44.46, 81.39)	38.89 (22.44, 58.25)	60.69 (47.75, 75.08)
INTP	46.78 (23.60, 64.23)	35.42 (12.90, 53.75)	34.52 (18.96, 49.89)
AFLP	47.51 (31.67, 63.52)	52.54 (24.75, 78.96)	46.43 (27.33, 68.22)

[a] Ninety-five percent confidence intervals are shown in parentheses.
[b] SCAP = South and Central American local pigs, MEDLP = Mediterranean and Slav local pigs, ANGLP = Anglo-Saxon local pigs, INTP = International pigs, and AFLP = African local pigs.
[c] RH = Roberts and Hiorns (1965) estimator, LC = Long (1991) and Chakraborty et al. (1992) estimator and W = Wang (2003) estimator.

(table 5). Most of genetic variation corresponded to the within-population variance component (σ^2_{WP}), particularly at the autosomal level ($\sigma^2 = 88.35\%$). Only 0–3% of the molecular variance was explained by the among-group variance component (σ^2_{AG}). This low value should be interpreted with caution because the two groups under consideration (pigs and wild boars) included both European and Far Eastern populations that are highly divergent at the genetic level (see for instance fig. 3), a feature that is expected to inflate the within groups variance. If we examine the coefficients of genetic differentiation among wild boar and pig populations, they are generally low (F_{ST} between European wild boars and Mediterranean pigs is 0.073 and between Far Eastern wild boars and pigs is 0.072) but highly significant ($P < 1.10^{-6}$) reflecting the existence of genetic differences between domesticated and wild *Sus scrofa*.

Population structure was somewhat higher when comparing pig and wild boar populations with different geographical origins (σ^2_{AGWP}: among-population and within-group variance component). Population structure was particularly strong for SSCY markers ($\sigma^2_{AGWP} = 35.26\%$ of total genetic variation) likely due to the differential geographical distribution of the HY3 haplotype. Mitochondrial and autosomal markers showed lower values for σ^2_{AGWP} (17.65% and 9.03%, respectively) consistent with the fact that several of the analyzed populations (e.g., African, International, and Anglo-Saxon pigs) have been formed by admixing European and Far Eastern breeds. Analysis of autosomal markers allowed us to detect large F_{ST} values between European wild boars versus Far Eastern pigs and wild boars ($F_{ST} = 0.19$ approximately). Genetic differentiation between European wild boars and International pig breeds was also considerably high ($F_{ST} = 0.14$), but this finding might be likely explained by the strong introgression of the latter population with Far Eastern breeds. Moreover, and as previously mentioned, all population comparisons yielded F_{ST} coefficients that were highly significant ($P < 1 \times 10^{-6}$) irrespective of their mean value. As a whole, these findings evidence the existence of a significant level of population structure in spite of the fact that most of genetic variability is distributed within rather than between populations.

Discussion

Weak Genetic Differentiation between Pigs and Wild Boars

We were interested in comparing the genetic diversity of pigs and wild boars, two populations that have undergone very distinct demographic processes, with the aim of defining to what extent domestication and selection have shaped the allelic pool of current swine breeds. According to our data, pigs and wild boars display similar levels of genetic variation. Nucleotide diversities for *MT-CYB* and SSCY markers were similar in pigs and wild boars. Microsatellite variation was also remarkably alike in wild and domesticated *Sus scrofa* (table 3). These results are consistent with those reported by Scandura et al. (2008) in European pigs and wild boars. In principle, domestication and breeding are expected to reduce genetic diversity because they involve the occurrence of founder effects, the utilization of a reduced number of sires with a high breeding value, and the progressive fixation of beneficial alleles through artificial selection (Innan and Kim 2004; Cruz et al. 2008). This would be particularly true for International pig breeds, such as Landrace and Large White, which have been strongly selected for increased growth, leanness, and prolificacy during the last century (Ojeda et al. 2008). There are several factors, however, that might explain why levels of genetic variation are similar in pigs and their wild ancestors. First, pig domestication rather than being a single event took place at multiple locations across Eurasia allowing the participation of *Sus scrofa* populations with different genetic backgrounds (Larson et al. 2005). Second, wild boar is a polygynous species that has undergone population bottlenecks as a consequence of excessive hunting (Scandura et al. 2008). Finally, and as we will discuss next, there might have been a sustained bidirectional gene flow between wild and domesticated *Sus scrofa* populations.

Analysis of autosomal genetic variation in pigs and wild boars showed that they are weakly although significantly differentiated (table 5). Moreover, pig and wild boar mitochondrial sequences clustered together in the median-joining network (fig. 1 and supplementary fig. 1, Supplementary Material online) and in the Bayesian phylogenetic tree (supplementary fig. 2, Supplementary Material online). In fact, the multidimensional scaling plot of F_{ST} coefficients (fig. 3) showed that the comparison of European versus Far Eastern *Sus scrofa* populations, instead of pigs versus wild boars, is much more meaningful, in terms of population structure. Consistently, Scandura et al. (2008) found that genetic divergence between European pigs and wild boars is very limited although significant. As mentioned above, this low genetic differentiation between pigs and their wild ancestors might be explained by the existence of a continuous bidirectional

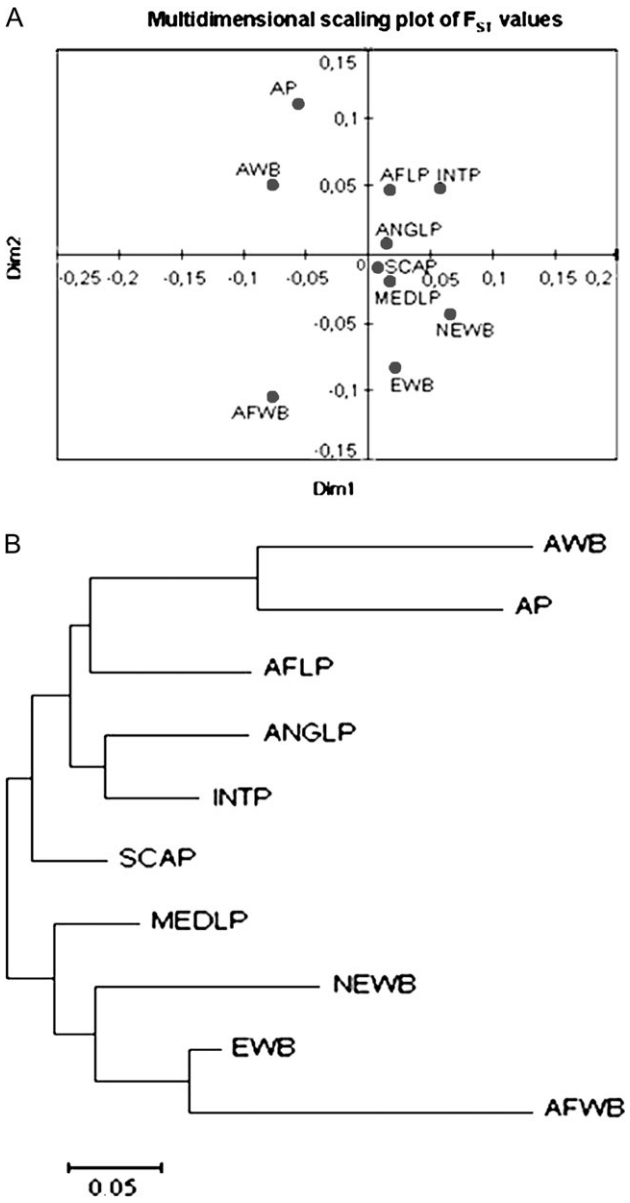

FIG. 3.—(*A*) Multidimensional scaling plot of F_{ST} values calculated with autosomal microsatellite data. All population comparisons yielded highly significant F_{ST} values ($P < 1\times 10^{-6}$) ranging from 0.03 to 0.19. Origin of each sample is indicated as follows: EWB = European wild boar, INTP = International pig, ANGLP = Anglo-Saxon local pig, MEDLP = Mediterranean and Slav local pig, AFWB = African wild boar, AFLP = African local, pigs, NEWB = Near Eastern wild boar, SCAP = South and Central American local pigs, AWB = Far Eastern wild boar and AP = Far Eastern local pigs. (*B*) Neighbor-Joining tree depicting the genetic relationships between the 10 aforementioned *Sus scrofa* populations. Microsatellite data were used to calculate D_A distances (Nei et al. 1983).

gene flow since their split 9,000 YBP. Conceivably, primitive pigs were raised in open herds, rather than being kept in small enclosures, giving a substantial chance for occasional interbreeding with their wild ancestor (Porter 1993). It is even possible that Neolithic farmers deliberately crossed domestic pigs with wild boars as a way to restock their herds (Vilà et al. 2005). This breeding strategy is even used nowadays for producing Iron Age pigs, which are a cross between Tamworth females and wild boar males (Porter 1993).

Origin of European Pig Breeds

The lack of a Near Eastern mitochondrial genetic signature in modern European pig breeds has been the main proof for supporting an independent domestication of pigs in Europe (Larson et al. 2005). Interestingly, mitochondrial analysis of ancient pig bones and teeth has evidenced that Near Eastern pigs entered Europe in the Neolithic but they did not leave any genetic signature in modern breeds (Larson et al. 2007). Mitochondrial data obtained in the current work show that Near Eastern and European wild boars share a common *MT-CYB* lineage named E1 (supplementary table 2, supplementary fig. 1, Supplementary Material online) but their gene pools are represented by distinct sets of *MT-CYB* haplotypes (sequences do not cluster together either in the median-joining network or in the Bayesian tree) revealing that they have independent origins. The E1 lineage, which in the study of Fang and Andersson (2006) is classified as European, can also be found in North Africa and Near East demonstrating that it has a much wider geographical distribution than previously anticipated. This lineage encompasses many different *MT-CYB* haplotypes of which the most frequent one is H_1 (GenBank accession number AY237496). This specific haplotype is shared by European, North African and, interestingly, two Armenian wild boars (fig. 1, supplementary table 2, Supplementary Material online). Larson et al. (2005) obtained similar results when analyzing the variability of the mitochondrial control region and reasoned that the presence of European haplotypes in Armenian wild boars might be due to introgression with European pig breeds However, this interpretation does not explain satisfactorily why H_1 is so frequent in wild boars from Tunisia and Morocco (mean frequency: 89%), two countries where pig farming is negligible. Moreover, if Near Eastern wild boars would have been significantly introgressed with modern International breeds, we would expect them to harbor Far Eastern *MT-CYB* haplotypes that, according to Fang and Andersson (2006), have a 29% frequency in International breeds. However, not even one of the 22 Near Eastern wild boars analyzed in the current work carried a Far Eastern *MT-CYB* haplotype. In the light of this reasoning, we favor a scenario where European, North African, and Near Eastern *Sus scrofa* share a certain number of *MT-CYB* lineages and haplotypes from ancient times and at variable frequencies (although the alternative hypothesis based on the occurrence of European introgressions cannot be completely ruled out).

Because mitochondrial markers only reflect the matrilineal history of a given species and they have a much faster lineage sorting and higher allelic extinction rate than autosomal markers (Zhang and Hewitt 2003), we have compared genetic variation of Near Eastern and European wild boar populations from a nuclear perspective. Analysis of microsatellite data showed that the F_{ST} value between European and Near Eastern wild boars is low (F_{ST} = 0.08) but significant. Moreover, STRUCTURE analysis evidenced that Near Eastern, North African, and European wild boars cluster together even when K is set to 8 (fig. 2), and this finding was additionally supported by a Neighbor-Joining tree based on the calculation of D_A distances (fig. 3*B*). SSCY haplotype frequencies were also very similar in

Table 5
AMOVA Results for Y-Chromosome, Autosomal, and Mitochondrial Markers Corresponding to Pigs and Wild Boars with a Worldwide Distribution

	Y-Chromosome Haplotypes				12 Autosomal Microsatellites				*MT-CYB* Haplotypes			
Source of Variation	df	Sum of Squares	Variance Components	% of Variation	df	Sum of Squares	Variance Components	% of Variation	df	Sum of Squares	Variance Components	% of Variation
Among groups[a]	1	1.653	−0.00439 Va	−2.42	1	76.320	0.12310 Va	2.62	1	3.063	−0.00617 Va	−1.40
Among populations within groups	8	13.240	0.06390 Vb	35.26	8	263.214	0.42486 Vb	9.03	8	22.358	0.07800 Vb	17.65
Within populations[b]	256	31.164	0.12173 Vc	67.17	726	3.018.387	4.15756 Vc	88.35	335	123.956	0.37002 Vc	83.74
Total	265	46.056	0.18123		735	3.357.921	470.552		344	149.377	0.44185	

[a] Groups ($N = 2$): Pigs versus wild boars.

[b] Populations under analysis ($N = 10$): EWB = European wild boar, INTP = International pigs, ANGLP = Anglo-Saxon local pigs, and MEDLP = Mediterranean and Slav local pigs, AFWB = African wild boar, AFLP = African local pigs, NEWB = Near Eastern wild boar, SCAP = South and Central American local pigs, and AWB = Far Eastern wild boar.

these three populations (table 2). Although a Near Eastern genetic signature has not been found in the mitochondrial genetic reservoir of current European pig breeds, a Near Eastern contribution to the nuclear gene pool cannot be ruled out. The recent advent of molecular tools enabling the high throughput genotyping of vast amounts of porcine autosomal single nucleotide polymorphisms should contribute to settle this issue.

An important landmark in the history of European pig breeds was the introgression of Chinese alleles into British swine populations during the 18th–19th centuries. As a consequence of this event, most International breeds (e.g., Large White, Landrace, and Pietrain) display Far Eastern mitochondrial haplotypes at variable frequencies, whereas local breeds with a more restricted geographical distribution (Iberian, Mangalitza, and others) do not (Giuffra et al. 2000; Clop et al. 2004; Fang and Andersson 2006). A novel conclusion that can be derived from our study is that this Chinese introgression of British breeds was fundamentally maternal because the HY3 SSCY haplotype, that is, relatively abundant in Far Eastern breeds, is completely absent from the International pig populations sampled in the current work. The consequences of this ancient introgression have been very significant from a genetic perspective because we have estimated that the proportion of Far Eastern alleles in the gene pools of Anglo-Saxon and International pig breeds are around 47–61% (table 4). Although these estimates have large confidence intervals and should be taken with caution, they show that these two populations were strongly introgressed with Far Eastern breeds. Similarly, Fang and Andersson (2006) have shown that the percentage of Far Eastern *MT-CYB* haplotypes in European swine has a mean value of 29% but the range of variation is very high depending on the breed and commercial line under consideration (Iberian: 0%, Mangalitza: 0%, Duroc 0%, Landrace: 0–43%, Large White: 14–100%, Piétrain: 0–78%, Berkshire: 78%, and Tamworth: 81%).

African Pig Breeds Display European and Far Eastern Genetic Signatures with a West versus East Geographical Distribution

The history of African pig breeds is largely unknown and controversial (Blench 1999). Archaeological evidence indicates that pigs were bred in the Ancient Egypt as soon as the fourth millennium and that they were widespread in North Africa. Probably, pigs reached Subsaharan Africa through the Nile corridor and they subsequently spread to West-Central Africa (Blench 1999). African native pigs were extensively admixed with exotic breeds as a result of the Portuguese exploratory journeys, beginning in the 15th century, as well as the colonization of the continent by several European countries in the 19–20th centuries (Blench 1999). By analyzing the genetic diversity of African pigs, we have evidenced the existence of a clear genetic dichotomy between East and West. The most frequent *MT-CYB* haplotype in pigs from Nigeria and Benin was H_1 (mean frequency: 50%), which is shared by European, North African and, to a lower extent, Near Eastern *Sus scrofa* populations. In the light of the current historical evidence (Blench 1999), the most plausible explanation for the presence of this haplotype in Subsaharan African pigs involves the occurrence of a European introduction (Africa was explored by the Portuguese in the 15th century), although a North African origin cannot be ruled out (European and North African wild boars are genetically similar so their genetic signatures are undistinguishable). More significantly, pigs from Nigeria and Benin did not display any of the *MT-CYB* (H_28, H_44, H_35) and SSCY (HY3) haplotypes that are characteristic of Far Eastern populations (supplementary tables 2, 4, and 5, supplementary fig. 1, Supplementary Material online). In strong contrast, pigs from Kenya and Zimbabwe carried Far Eastern *MT-CYB* (e.g., H_28, H_44, and H_50) and SSCY (e.g., HY3 is fixed in Mukota pigs) haplotypes at high frequencies (supplementary tables 2, 4, and 5, supplementary fig. 1, Supplementary Material online). In summary, a clear Far Eastern genetic signature can be found in pigs from the Indian Ocean coast of Africa but not in their counterparts of the Atlantic seaboard. Two alternative scenarios might explain the presence of Far Eastern alleles in East African breeds, that is, they might have been introduced through either a European intermediary, given that British breeds were strongly admixed with Chinese pigs in the 18–19th centuries (Porter 1993), or by direct introgression with Far Eastern breeds. In this sense, the multidimensional scaling plot of F_{ST} values and the Neighbor-Joining tree based on D_A distances evidences that African pigs occupy an intermediate position between Far Eastern pigs and wild boars and Anglo-Saxon and International swine breeds,

making it difficult to discern which of the two scenarios is correct. These findings are probably due to the fact that 79% of African pigs analyzed with microsatellites are originally from East Africa, where the Far Eastern introgression has been particularly strong. Similarly, the STRUCTURE analysis suggests a close relatedness between African pigs and International breeds. For instance, with a number of clusters $K = 3$ or 4, both populations appear to share a common genetic background ($K = 3$, in green; $K = 4$, in yellow). However, the analysis of Y-chromosome genetic diversity shows that HY3, an SSCY haplotype completely absent from International breeds, has an average frequency of 0.53 in Kenyan and Mukota pigs. This finding gives strong support to the hypothesis of a direct Far Eastern introduction in Africa. The presence of Chinese-style lard pigs in Zimbabwe and Mozambique is very consistent with this interpretation (Blench 1999).

The precise geographical origin of this Far Eastern introduction is not clear because African *MT-CYB* haplotypes clustered indistinctly with sequences from Chinese, Indonesian, Korean, Sinhalese, and Vietnamese pigs (supplementary fig. 3, Supplementary Material online). The time at which this introduction took place is also hard to infer. Throughout the ages, Eastern Africa has been a major point of entry of several livestock species into the continent including chicken, camel and zebu (Hanotte et al. 2002; Muchadeyi et al. 2008). Ancient contacts, through the Indian Ocean, between Far East and Africa, have been widely documented. For instance, Madagascar was settled by Indonesian seafarers 1,500–2,000 YBP (Hurles et al. 2005) and an active trade between China and Somalia is reported in the Duan Chengshi writing dating to AD 863 (Levathes 1994). Moreover, the Portuguese might have transported pigs from Macau to Zimbabwe and Mozambique a few centuries ago, giving origin to African breeds with an Asian-like phenotype (Blench 1999). In summary, the introduction of Far Eastern pigs in Africa might be likely explained by the existence of ancient commercial routes linking Asia and Africa through the Indian Ocean.

The Origin of South American Pig Breeds Is Fundamentally European

Pigs were first brought to South America by the Spanish and Portuguese colonizers at the end of the 15th century (Rodero et al. 1992; Delgado et al. 2004). These pigs had Iberian, Celtic, and Canarian origins and showed an excellent adaptation to this new environment (Rodero et al. 1999; Delgado et al. 2004). The Portuguese colonizers might have subsequently admixed this founder population with pigs from China, Vietnam, and Thailand (Porter 1993). This Asian heritage is still evident in several Brazilian breeds such as Canastrinho and Nilo (Porter 1993). Besides, South American breeds have been extensively hybridized with International breeds such as Duroc and Poland China (Porter 1993). Our genetic analysis of South American swine populations showed that the two most abundant *MT-CYB* haplotypes are H_1 and H_11 (supplementary table 2, supplementary fig. 1, Supplementary Material online). Current historical evidence strongly suggests that H_1 might have been introduced by the European colonizers several centuries ago (Rodero et al. 1999; Delgado et al. 2004). This interpretation is supported by the fact that H_11 is shared by South American and Mediterranean pigs (supplementary table 2, supplementary fig. 1, Supplementary Material online). Moreover, the multidimensional scaling plot of F_{ST} values inferred from microsatellite allele frequencies showed that South American and Mediterranean pigs are closely related (fig. 3*A*). Similarly, Souza et al. (2009) have recently demonstrated that several Brazilian breeds have an Iberian origin. These findings are clearly consistent with the Spanish and Portuguese ancestry of South American creole populations (Rodero et al. 1999). Interestingly, we have detected the presence of Far Eastern *MT-CYB* haplotypes (H_8, H_9, and H_10) in Nicaraguan pigs. Consistently, we have observed the segregation of the HY3 haplotype in a few Nicaraguan and Argentinian pigs (table 2, supplementary tables 4 and 5, Supplementary Material online). As reasoned above, the null frequency of this SSCY haplotype in International pig breeds suggests that Far Eastern breeds might have participated to some extent in the foundation of American creole populations. Recent *MT-CYB* data showing that several Brazilian local breeds and commercial lines harbor Far Eastern haplotypes gives strong support to this hypothesis (Souza et al. 2009). The existence of important Portuguese (from Macau to Brazil) and Spanish (from Philippines to Mexico) commercial routes connecting Europe and Asia through America might provide a historical framework for these findings.

Final Conclusions

The independent domestication of pigs in Europe and Far East generated two highly differentiated gene pools that spread to Africa and South America, likely following exploratory and commercial sea routes that were specially active during the European colonization of these two continents. British pig breeds were also strongly introgressed with Chinese sows in the 18th–19th centuries with the aim of improving fatness and reproductive traits. This worldwide process of population admixture combined with the occurrence of a sustained gene flow between pigs and wild boars might have played a major role in preserving the genetic variability of current porcine breeds.

Supplementary Material

Supplementary tables 1–5 and supplementary figures 1 and 2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We are grateful to Carles Vilà, Carles Lalueza, and Julio Rozas for critically reading the manuscript. We are also indebted to Maria Saña for valuable comments about the process of pig domestication and to Francesc Calafell for helping in the statistical analyses. Many thanks to all of those contributing to sample collection: G. Aktan (Ay-Fi

Safari), A.R. Jafari-Tabrizi (Iran Safari), S. Shushunov (Russian Hunting Agency), F. Pilla (University of Molise), J. Tibau (Institut de Recerca i Tecnologia Agroalimentàries), J. Capote (Instituto Canario de Investigaciones Agrarias), A. Linden (University of Liege), J.L. Noguera (Institut de Recerca i Tecnologia Agroalimentàries), L. Silió and M.C. Valdovinos (Instituto Nacional de Investigaciones Agrarias, Spain), E. Martínez (Zoo of Madrid), N. Okumura (STAFF Institute Japan), G. Rohrer (United States Department of Agriculture), Z. Bosze (Agricultural Biotechnology Center, Hungary), M. Cumbreras (Diputación de Huelva, Spain), E. von Eckardt (Sweden), M.T. Paramio (Universitat Autònoma de Barcelona, Spain), O. Vidal (Universitat de Girona, Spain), M. Kovac (Slovenia), D. Vidal and C. Gortázar (Instituto de Investigación en Recursos Cinegéticos, Spain), Instituto de Capacitación del Oriente (Bolivia), Veterinarios Sin Fronteras (Spain). A.O. and D.G. were recipients of PhD grants from the Ministerio de Educación y Ciencia, (Spain) and Universitat Autònoma de Barcelona. O.R. received funding from the Molecular Genetics Veterinary Service of the Universitat Autònoma of Barcelona (http://svgm.uab.es/svgm). This study was partially funded by projects AGL2002-04271-C03-03 and AGL2007-66707-C02-02 (Ministry of Education and Science, Spain) to M.A.; AGL2004-0103, AECI/5267/06, and BFU2004-02253 (Ministry of Education and Science, Spain) to M.P.E.; AGL2000-1229-C03-03 (Ministry of Education and Science, Spain) to A.S. and National Natural Science Foundation of China (grant 30425045) to L.S.H.

Literature Cited

Alves E, Ovilo C, Rodríguez MC, Silió L. 2003. Mitochondrial DNA sequence variation and phylogenetic relationships among Iberian pigs and other domestic and wild pig populations. Anim Genet. 34:319–324.

Bandelt HJ, Macaulay V, Richards M. 2000. Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. Mol Phylogenet Evol. 16:8–28.

Blench . 1999. A history of pigs in Africa. In: Blench RM, Mac Donald K, editors. Origins and development of African livestock: archaeology, genetics, linguistics and ethnography. Florence, KY: Routledge Books. p. 355–367.

Chakraborty R RM, Kamboh MI, Nwankwo M, Ferrell RE. 1992. Caucasian genes in American blacks: new data. Am J Hum Genet. 50:145–155.

Chevenet F, Brun C, Bañuls AL, Jacq B, Christen R. 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics. 7:439.

Clop A, Amills M, Noguera JL, Fernández A, Capote J, Ramón MM, Kelly L, Kijas JM, Andersson L, Sànchez A. 2004. Estimating the frequency of Asian cytochrome B haplotypes in standard European and local Spanish pig breeds. Genet Sel Evol. 36:97–104.

Cruz F, Vilà C, Webster MT. 2008. The legacy of domestication: accumulation of deleterious mutations in the dog genome. Mol Biol Evol. 25:2331–2336.

Delgado JV, Camacho ME, León JM, de la Haba MR, Vallecillo A, Barba C, Cabello A. 2004. Poblaciones porcinas de Iberoamérica. In: Delgado JV, editor. Biodiversidad porcina Iberoamericana. Servicio de Publicaciones de la Universidad de Córdoba. Córdoba (Spain). p. 21–32.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software Structure: a simulation study. Mol Ecol. 14:2611–2620.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.1: an integrated software package for population genetics data analysis. Evol Bioinform Online. 1:47–50.

Fang M, Andersson L. 2006. Mitochondrial diversity in European and Chinese pigs is consistent with population expansions that occurred prior to domestication. Proc Biol Sci. 273:1803–1810.

Fang M, Hu X, Jiang T, Braunschweig M, Hu L, Du Z, Feng J, Zhang Q, Wu C, Li N. 2005. The phylogeny of Chinese indigenous pig breeds inferred from microsatellite markers. Anim Genet. 36:7–13.

Ferro M. 1997. History of colonisation. Oxford: Routledge.

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

Hanotte O, Bradley DG, Ochieng JW, Verjee Y, Hill EW, Rege JE. 2002. African pastoralism: genetic imprints of origins and migrations. Science. 296:236–239.

Hasegawa M, Kishino H, Yano T. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol. 22:160–174.

Hurles ME, Sykes BC, Jobling MA, Forster P. 2005. The dual origin of the Malagasy in Island Southeast Asia and East Africa: evidence from maternal and paternal lineages. Am J Hum Genet. 76:894–901.

Innan H, Kim Y. 2004. Pattern of polymorphism after strong artificial selection in a domestication event. Proc Natl Acad Sci USA. 101:10667–10672.

Kim KI, Lee JH, Li K, Zhang YP, Lee SS, Gongora J, Moran C. 2002. Phylogenetic relationships of Asian and European pig breeds determined by mitochondrial DNA D-loop sequence polymorphism. Anim Genet. 33:19–25.

Kumar S, Subramanian S. 2002. Mutation rates in mammalian genomes. Proc Natl Acad Sci USA. 99:803–808.

Langella O. 2002. Population 1.2.28. Logiciel de génétique des populations. CNRS UPR 9034, Gif-sur.Yvette (France): Laboratoire Populations, Génétique et Evolution. http://www.cnrs-gif.fr/pge

Larson G, Albarella U, Dobney K, et al. (20 co-authors). 2007. Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. Proc Natl Acad Sci USA. 104:15276–15281.

Larson G, Dobney K, Albarella U, et al. (13 co-authors). 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science. 307:1618–1621.

Levathes L. 1994. When China ruled the seas. New York: Simon & Schuster. ISBN 0-671-70158-4.

Long JC. 1991. The genetic structure of admixed populations. Genetics. 127:417–428.

Megens HJ, Crooijmans RP, San Cristobal M, Hui X, Li N, Groenen MA. 2008. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. Genet Sel Evol. 40:103–128.

Muchadeyi FC, Eding H, Simianer H, Wollny CB, Groeneveld E, Weigend S. 2008. Mitochondrial DNA D-loop sequences suggest a Southeast Asian and Indian origin of Zimbabwean village chickens. Anim Genet. 39:615–622.

Nei M, Tajima F, Tateno T. 1983. Accuracy of estimated phylogenetic trees from molecular data. J Mol Evol. 19:153–170.

Ojeda A, Wang L, Ren J, Angiolillo A, Cho IC, Soto H, Lemús-Flores C, Makuza SM, Folch JM, Pérez-Enciso M. 2008. Selection in the making: a worldwide survey of haplotypic diversity around a causative mutation in porcine *IGF2*. Genetics. 178:1639–1652.

Page RD. 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci. 12:357–358.

Porter V. 1993. Pigs. A handbook to the breeds of the world. Mountfield (United Kingdom): Helm Information Ltd.

Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics. 14:817–818.

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

Roberts DF, Hiorns RW. 1965. Methods of analysis of the genetic composition of a hybrid population. Hum Biol. 37:38–43.

Rodero A, Delgado JV, Rodero E. 1992. Primitive Andalusian livestock and their implications in the discovery of America. Arch Zootec. 41:383–400.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19:1572–1574.

Rousset F, Raymond M. 1995. Testing heterozygote excess and deficiency. Genetics. 140:1413–1419.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 19:2496–2497.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406–425.

SanCristobal M, Chevalet C, Haley CS, et al. (32 co-authors). 2006. Genetic diversity within and between European pig breeds using microsatellite markers. Anim Genet. 37:189–198.

Scandura M, Iacolina L, Crestanello B, Pecchioli E, Di Benedetto MF, Russo V, Davoli R, Apollonio M, Bertorelle G. 2008. Ancient *vs*. recent processes as factors shaping the genetic variation of the European wild boar: are the effects of the last glaciation still detectable? Mol Ecol. 17:1745–1762.

Souza , Paiva SRCA, Pereira RW, Guimarães SE, Dutra Jr WM, Murata LS, Mariante AS. 2009. Iberian origin of Brazilian local pig breeds based on Cytochrome b (*MT-CYB*) sequence. Anim Genet. Forthcoming.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Schaffner SF. 2004. The X chromosome in population genetics. Nat Rev Genet. 5:43–51.

Vilà C, Seddon J, Ellegren H. 2005. Genes of domestic mammals augmented by backcrossing with wild ancestors. Trends Genet. 21:214–218.

Wang J. 2003. Maximum-likelihood estimation of admixture proportions from genetic data. Genetics. 164: 747–765.

Wu GS, Yao YG, Qu KX, Ding ZL, Li H, Palanichamy MG, Duan ZY, Li N, Chen YS, Zhang YP. 2007. Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. Genome Biol. 8:R245.

Zhang DX, Hewitt GM. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Mol Ecol. 12:563–584.

Anne Stone, Associate Editor

Accepted June 1, 2009