# The Origin of the Domestic Pig: Independent Domestication and Subsequent Introgression

# E. Giuffra,<sup>1</sup> J. M. H. Kijas,<sup>1</sup> V. Amarger, Ö. Carlborg, J. T. Jeon and L. Andersson

Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala Biomedical Centre, S-751 24 Uppsala, Sweden

> Manuscript received June 18, 1999 Accepted for publication November 29, 1999

### ABSTRACT

The domestic pig originates from the Eurasian wild boar (*Sus scrofa*). We have sequenced mitochondrial DNA and nuclear genes from wild and domestic pigs from Asia and Europe. Clear evidence was obtained for domestication to have occurred independently from wild boar subspecies in Europe and Asia. The time since divergence of the ancestral forms was estimated at  $\sim$ 500,000 years, well before domestication  $\sim$ 9,000 years ago. Historical records indicate that Asian pigs were introduced into Europe during the 18th and early 19th centuries. We found molecular evidence for this introgression and the data indicated a hybrid origin of some major "European" pig breeds. The study is an advance in pig genetics and has important implications for the maintenance and utilization of genetic diversity in this livestock species.

THE wild boar is widespread in Eurasia and occurs I in Northwest Africa; the existence of at least 16 different subspecies has been proposed (Ruvinsky and Rothschild 1998). Domestication of the pig is likely to have occurred first in the Near East  $\sim$ 9000 YBP and may have occurred repeatedly from local populations of wild boars (Bökönyi 1974). However, it is not yet established whether modern domestic pigs showing marked morphological differences compared with their wild ancestor have a single or multiple origin. Darwin (1868) recognized two major forms of domestic pigs, a European (Sus scrofa) and an Asian form (Sus indicus). The former was assumed to originate from the European wild boar, while the wild ancestor of the latter was unknown. Darwin considered the two forms as distinct species on the basis of profound phenotypic differences. It is well documented that Asian pigs were used to improve European pig breeds during the 18th and early 19th centuries (Darwin 1868; Jones 1998) but to what extent Asian pigs have contributed genetically to different European pig breeds is unknown. In a recent study the divergence between major European breeds and the Chinese Meishan breed was estimated at  ${\sim}2000$ years using microsatellite markers (Paszek et al. 1998). Limited studies of mitochondrial DNA (mtDNA) have indicated genetic differences between European and Asian pigs but no estimate of the time since divergence has been provided (Watanabe et al. 1986; Okumura et al. 1996).

The aim of this investigation was to provide a more comprehensive molecular analysis of the origin of domestic pigs by analyzing wild and domestic pigs from both Asia and Europe. The molecular analysis included the entire coding sequence of the mtDNA cytochrome B (cvtB) gene and 440 bp of the hypervariable mtDNA control region. In addition, we have analyzed three nuclear genes: melanocortin receptor 1 (MC1R), tyrosinase (TYR), and a glucosephosphate isomerase pseudogene (GPIP). We have reported previously that MC1R mutations cause coat color variation in pigs (Kijas et al. 1998). An interesting finding was that the allele for black color in Asian Meishan and British Large Black pigs differed by as many as four nucleotide substitutions from the consensus of MC1R alleles in European domestic and wild pigs. The observation was interpreted to reflect the introgression of an MC1R allele of Asian origin to the Large Black population. *TYR* encodes the tyrosinase enzyme, which has a key role in pigment synthesis. Loss-of-function mutations in this gene cause albinism in many species. We investigated this gene previously in a wild boar/Large White intercross but the observed sequence polymorphism did not correlate with any coat color variation. GPIP is clearly a pseudogene as it contains many potentially inactivating mutations (Harbitz et al. 1993). The gene was included in this study because we assume that the selection pressure on this sequence should be minimal.

# MATERIALS AND METHODS

Corresponding author: L. Andersson, Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala Biomedical Centre, Box 597, S-751 24 Uppsala, Sweden. E-mail: leif.andersson@bmc.uu.se Animals: Hair or bl domestic and wild bo Large White, Landrace all originated from Swe

**Animals:** Hair or blood samples were available from the domestic and wild boar populations listed in Table 1. The Large White, Landrace, Hampshire, and Duroc domestic pigs all originated from Swedish populations. The Meishan domestic pigs were from the PiGMaP reference pedigrees for gene

<sup>&</sup>lt;sup>1</sup>These authors contributed equally to this work.

## **TABLE 1**

		Cytochrome B variants											
Population	п	A1	A2	A3	E1	E2	E3	E4					
Asia													
Wild boar, Japan	7	6	_	1	_	_	_	_					
Meishan	7	7	_	_	_	_	_	_					
Europe													
Wild boar, Italy	24	_	_	_	23	_	1	_					
Wild boar, Poland	15	_	_	_	12	3	_	_					
Large White	27	11	2	_	13	_	_	1					
Landrace	13	3	1	_	9	_	_	_					
Hampshire <sup>a</sup>	20	_	_	_	20	_	_	_					
Duroc <sup>a</sup>	12	2	_	_	10	_	_	_					
Mangalica	1	_	_	_	1	_	_	_					
Other areas													
Wild boar, Israel	3	_	_	_	3	_	_	_					
Domestic pig, Cook Island	1	_	_	_	1	_	_	_					

Distribution of mitochondrial cytochrome B variants among domestic pig and wild boar populations

The different variants were distinguished by SSCP. The corresponding sequences are defined in Figure 1B. <sup>a</sup> Hampshire and Duroc are European-type domestic pigs but were developed mainly in the United States.

mapping (Archibal d *et al.* 1995). Genomic DNA was extracted from blood by standard methods or from hair by chelex extraction.

Analysis of mitochondrial DNA: The primers LmPro/L15997 and TDKT/H16498 (Ward *et al.* 1991) were used to amplify and sequence a 440-bp fragment that included  $\sim$ 20 bp of the tRNA<sup>Pro</sup> and 420 bp of the hypervariable (5') portion of the control region. The complete (1140-bp) cytochrome B sequence was amplified by PCR and was sequenced by using combinations of the primers L14724, L14841, L15408, H15149, H15915 (Irwin *et al.* 1991) and the pig-specific primers L0616 (5'-TATTCCTGCACGAAACCGGAT) and H1106 (5'-AGGTTGTTTTCGATGATGCTAG). PCR products were sequenced directly using the Big Dye terminator chemistry and an ABI377 Prism DNA sequencer (Perkin Elmer, Norwalk, CT).

Sequence analysis was carried out on a subset of animals while a screening method based on single strand conformation polymorphism (SSCP) was developed for testing all animals included in the study. A 377-bp PCR product from *cytB* obtained with the primers L14841 and H15149 was subjected to SSCP analysis using 10% nondenaturating polyacrylamide gels as previously described (Marklund *et al.* 1995).

The GenBank accession numbers for the mitochondrial sequences described in this work are AF136555-AF136568, AF182446 (control region); AF136541-AF136554, AF163099, AF163100 (*cytB*).

**Detection of sequence polymorphism in three nuclear genes:** The primers and reaction conditions used to amplify and sequence *MC1R* have been described (Kijas *et al.* 1998). A 727-bp region (including primers) of *TYR* exon 1 was amplified with the primers TYRPe1F (5'-AGGGGTAGCTGGAAAGA GAA-3') and TYRPe1R (5'-CAATACCAGCAAGAAGAGGTC-3') using standard conditions but with the addition of 5% DMSO. SSCP analysis of *TYR* was carried out with the internal primers Tyrex1JK1 (5'-ACTACCAGCCCAGACTTAGTC-3') and Tyrex1JK2 (5'-AGCAAAATCAATGTCTCTCCAG-3'), which were used to amplify a 173-bp product spanning polymorphic positions 544 and 628 (see Figure 3). Denaturated PCR product (1 µl) solution was loaded on gels and was electrophoresed for 17 hr at 1.0 W in a 10% native polyacrylamide gel (130 × 130 × 0.5 mm) before visualization using standard silver staining. The primers GPIP1 (5'-TGCAGTTGAGAAGGACTTTA CTT-3') and GPIP4 (5'-GTATCCCAGATGATGTCATGAAT-3') were used to specifically amplify 784 bp of *GPIP*. The primers included eight and seven mismatched positions, respectively, compared with the functional *GPI* gene (Harbitz *et al.* 1993). Internal primers GPIP2 (5'-CTTGCCCTGC TGGCTCTGCC-3') and GPIP3(5'-ATCCATCCTTACCAGA TGCTGG-3') amplified a 308-bp fragment (spanning polymorphic positions 316, 388, 389, and 415; see Figure 3) that was used for SSCP analysis as described above.

The GenBank accession numbers for the nuclear sequences described in this work are AF181958-AF181961 (*GPIP*); AF181962, AF181963 (*TYR*); AF181964 (*MC1R*).

Sequence analysis: Some of the sequences used for mitochondrial DNA phylogenetic analysis were already described in the database (GenBank accession nos.: AB015067, AB01-5069-AB015072, AB015074, AB015075, AB015077, AB015079, AB015080, AB015082, AB015083, AB015085-AB015090, AB015094, AB015095, D42171). Sequence divergence (K) and standard error calculations for TYR and GPIP sequences were calculated using the MEGA software (Kumar et al. 1993). Time of divergence was estimated using the simple equation T =K/2r, where r is the rate of nucleotide substitutions (Li 1997). Genetic distances between mitochondrial sequences were corrected for multiple hits with Kimura's two-parameter method (Kimura 1980). Sites representing gaps in the aligned sequence were treated as missing. Neighbor-joining trees were constructed using the PAUP 4.0 software (Swofford 1998). Bootstrap analyses (using 1000 replications) were used to assess the confidence in branching order. Maximum parsimony trees were constructed using PAUP 4.0 and the branch-andbound method. The average pairwise distance between mtDNA cytB sequences from the European and Asian clades was calculated according to Nei (1987).

#### RESULTS

Sequence analyses reveal three distinct pig mtDNA clades: We sequenced the entire coding sequence (1140

bp) of the mtDNA *cytB* gene and 440 bp of the mtDNA control region from European wild boars from Poland and Italy, wild boars from Israel, Asian wild boars from Japan, European domestic pigs, Chinese Meishan domestic pigs, and one pig from the Cook Islands (Table 1). Some sequence information was also available in the nucleotide database. The phylogenetic analysis of the hypervariable D-loop region displayed three distinct mtDNA clades, one Asian and two European (Figures 1A and 2A). The Asian clade included Japanese wild boars, Chinese Meishan pigs, and some European domestic pigs. European clade 1 was composed of the majority of European wild boars, Israeli wild boars, most European domestic pigs, and the pig from Cook Island.

Α Nucleotide positions 1 5 7 5 5 55 55 5 7 5 7 5 7 5 7 5 7 5 7 5 5 5 6 6 66 66 6 8 5 5 8 5 4 4 5 6 6 7 8 8 9 1 1 2 8 9 1 1 2 з з 4 5 2 2 Clade Animals 0 9 6 з GTGAAGC - CCATTTTTGCTCACACCCT EI EWB2 13 Cook т -. . . . . c EWB1 EL A . . . . . . . . . . . . . . Т С . С . A . ΕL L2 . . . . . . . H1, IWB1-3 . . • ΕI . . . . . . . . . . ΕI H2 А . . . . . ΕI Ma - . . С . ΕI L1 . . . . ÷ . • . . LW2 А . . ΕI • - . T - . T ΕII EWB3 . G. . G . • · . · : С т G А . С EWB4 G . . . . . т G ΕII А . · G., . . . . т EWB5 Α . G G . . . . . С G ΕII . -т С . AWB6 С т G . . т G Α . • А . . . . . . AWB7 С . . . А - - т т G . . . . . т G A . . . - - т A AWB8, 9 С . . . А т G . . С . . . . т G T A AWB10 с . . . Α - - тт G С . . . . . . G G т A - - T A - - T A - - T A - - T A - - T A - - T A AWB11 С . . . Т G С . . . . G т T G C T G C T G C T G C T G C A **AWB12** С . . . . . . G . . ΤG . c т . . . . . A AWB13 С . . . . . . AWB14, LW1 · · ΤG Т т A С · . . . c GCC G т т т С . A Me . . . . .

ΤG

A D

в														Ν	uci	leo	tid	e p	osi	itio	ns															
_		*		*													*													٠	*					
		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5
		1	2	2	3	3	3	3	з	4	4	4	5	5	5	7	7	7	7	8	8	8	9	0	0	0	0	0	0	0	1	1	2	2	2	2
		6	0	4	0	6	7	8	9	0	2	6	2	8	9	0	1	8	9	0	5	8	3	0	2	2	з	з	4	4	0	1	0	1	3	5
SSCP	Animals	7	0	4	9	9	8	4	6	5	9	5	2	8	4	5	3	0	5	7	5	8	0	2	0	9	6	8	1	5	2	3	0	2	6	7
	EWB1, L1, H1,																																			
E1	IWB1-3, Cook	С	т	Т	G	Т	С	T	т	т	А	т	А	Α	G	С	Т	С	С	А	Α	Т	С	Т	т	т	т	G	С	G	G	т	G	С	С	G
E2	EWB2								С				•	•		•		·	•	•	•		•	·	·	•	·	٠	·	·	•	·	•	·	•	·
E1	L2								•					•		•		•		·	•	·	•	٠	·	С	·	•	•	•	•	٠	•	÷	٠	·
E1	H2									•			G			٠	•	•	•		•	•	·	٠	•	·	·	•	٠	·	•	٠	•		·	•
E1	Ma													•				•	Т		•	•			С	•	•		•	·	·	·	·	٠	·	Α
E4	LW2						•	С		•					•		•				•		•		•	•	٠	•	•	·		•	·	٠	·	Α
E3	EWB3				Α	С				С				-	•	•	•	т	•	•				•	•	•	•	·	Т	·	Α	•	·	·	•	Α
-	EWB4	Т		С	Α	С				С		•				т		Т	•		•		т	•	•	•	•	·	Т	•	Α	С	·	•	·	•
-	EWB5	т		С	Α	С				С						т	•	т				•	т	•			-	•	т	·	Α	•	·	٠	•	•
-	AWB6, 8					С				С		-		•		Т		т			•	С	•	•				А	Т	•	Α	•	А	•	т	•
-	AWB7					С				С						т	С	т		•	•	С	•		•		•	А	т	•	А	•	Α	•	т	•
-	AWB9					С				С						т	¢	т	-		•	С		•	•			А	Т	•	А	•	•	•	Т	•
A 1	AWB10									С	G	С				т					G			С			С	А	т		Α	•	А	•	Т	•
A2	LW1					С	т			С	G	С			А	т					G			С			С	А	т		Α	•	Α	т	т	٠
A 1	Me, D					С		-		С	G	С				т					G			¢			С	А	т	А	Α		А	·	T	Α
A3	AWB11									С	G			G		т				т	G			С			•	А	т	•	Α	С	А	•	т	•
-	AWB12					С				С	G					т					G			С				А	т		Α	С	А	•	Т	•
-	AWB13, 14		С	•	•	С		•		С	G	·	•	·	-	Т	•	·	•	•	G	·	·	С	·	·	·	Α	т	·	Α	•	A	·	т	•

с

The European clade 2 included only three wild boars from Southern Europe (Italy). All mtDNA haplotypes having a presumed European origin and found in domestic pigs belonged to European clade 1. The phylogenetic analysis of the cytochrome B sequences (Figures 1B and 2B) displayed again a statistically supported Asian clade but the presence of two distinct European clades was not statistically significant. European and Asian haplotypes could be distinguished easily by SSCP analysis of a 377-bp PCR product from the *cytB* gene. SSCP screening confirmed the conclusions from the tree analysis and showed that Asian mtDNA haplotypes are present in European Large White, Landrace, and Duroc domestic pigs (Table 1). Very similar tree topolo-

> Figure 1.—(A) Variable positions in the 5' region of the mitochondrial DNA control region (440 bp) among wild boars and domestic pigs. Sequence identities and deletions are indicated by dots and dashes, respectively; missing data are indicated by blanks. Nucleotide positions are numbered according to the complete pig mitochondrial DNA reference (Ursing and Arnason 1998). The first column indicates assignments of haplotypes to phylogenetic clades. Abbreviations and geographical location for the wild boar samples are as follows: EWB, European wild boar; IWB. Israeli wild boar; AWB, Asian wild boar; EWB1-2, Poland; EWB3-5, Italy; AWB6-9, Ryukyu Island, Southern Japan; AWB10-14, Japan. Abbreviations for domestic pig breeds are as follows: L, Swedish Landrace; H, Hampshire; Ma, Mangalica; LW, Large White; Me, Meishan; D, Duroc; Cook, Cook Island. (B) Variable positions in the mitochondrial cytochrome B gene (1140 bp) among wild boars and domestic pigs. The six positions involving amino acid substitutions are marked by asterisks. The first column indicates assignments of the haplotypes according to the SSCP analysis; dashes refer to the sequences available only in the nucleotide database that were not subjected to SSCP analysis (Table 1).



Figure 2.—Neighbor-joining trees of wild boar and domestic pig mitochondrial DNA haplotypes based on (A) 440 bp of the control region and (B) 1140 bp of the *cytB* gene. Bootstrap values (threshold of 50% after 1000 replicates) are reported on the nodes. All abbreviations are explained in the legend to Figure 1.

gies for both the control region and *cytB* were obtained with maximum parsimony and maximum likelihood analysis (data not shown).

The fact that mtDNA sequences from some domestic pigs are closely related to European wild boar sequences, whereas others cluster with Asian wild boar sequences, provides conclusive evidence for independent domestication of pigs in Europe and Asia. A sequence divergence of ~2% per million years for mtDNA has been found among mammalian species, corresponding to a nucleotide substitution rate of  $1 \times 10^{-8}$  substitutions/site/year in each lineage (Brown *et al.* 1979). The

average pairwise nucleotide distance between European and Asian mtDNA cytB sequences belonging to clade E1 and A was  $17 \pm 1.4$  substitutions  $(1.45 \pm 0.12\%)$ . A lower estimate of 1.0% was obtained by calculating the interpopulation distance, which attempts to take into account the genetic diversity existing in the common ancestor of the two populations (Nei 1987). This more conservative estimate indicates that the time since divergence of the two populations is of the order of 500,000 YBP, evidently well before domestication  $\sim$ 9000 YBP. The most likely explanation for the fairly high frequency of Asian mtDNA haplotypes in some European domestic breeds is that this is due to the documented introgression of Asian pigs during the 18th and early 19th centuries. The maternal inheritance of mtDNA implies that Asian sows were used for introgression and this is consistent with both written records and contemporary art.

Sequence polymorphism in three nuclear genes is consistent with a European and Asian origin of domestic pigs: The genetic diversities of three nuclear genes, MC1R, TYR, and GPIP, were also investigated. Our previous MC1R study (Kijas et al. 1998) showed that the MC1R\*2/ dominant black allele present in Chinese Meishan and British Large Black pigs shares two synonymous and two nonsynonymous substitutions absent from MC1R alleles in other European domestic breeds and in European wild boars (MC1R\*1-\*4; Figure 3A). The observation that the MC1R sequence from Japanese wild boars (MC1R\*5; Figure 3A) shares one of the synonymous substitutions found in MC1R\*2 indicates that the latter originates from an Asian rather than a European ancestor. MC1R is essentially monomorphic within domestic breeds due to strong selection for coat color homogeneity and we have not observed any heterogeneity among wild boars. We did not attempt to use MC1R for calculating time since divergences because the observed rate of nucleotide substitutions may be affected by the strong phenotypic selection at this locus.

We sequenced the major part of *TYR* exon 1 (727 bp) from two animals each of European and Japanese wild boars and several domestic breeds. Two alleles differing by four synonymous substitutions were found (Figure 3B). There were no fixed differences between continents but *TYR\*1* occurred predominantly in Japanese wild boars and Meishan domestic pigs, while *TYR\*2* was most common in European wild boars and domestic pigs (Table 2).

Since both *MC1R* and *TYR* are coding sequences, the *GPIP* pseudogene was included as a noncoding nuclear sequence. Sequence analysis of 784 bp of this pseudogene from 14 animals representing domestic and wild pigs from both continents revealed four alleles differing by multiple nucleotide substitutions (Figure 3C). Once again, there was a clear tendency for more pronounced allele frequency differences between continents than between wild and domestic pigs within continents,

А	Position 2 2 3 3 4 7 7 7 9 5 6 8 1 2	В	Position 356 3342	с	Position 1 2 3 3 3 4 5 7 8 2 1 8 8 1 4 6	Figure 3.—Sequence polymorphism in three nu-
Allele	4641280	Allele	1148	Allele	15689541	clear genes in the pig: (A) MC1R, (B) TYR exon 1, and
MC1R*1 MC1R*2 MC1R*3 MC1R*4 MC1R*5	G T T G C G G A C C A A T A -	TYR*1 TYR*2	G C T T A T C A	GPIP*1 GPIP*2 GPIP*3 GPIP*4	G G A G A C C G A T T A T T - T A A G C T T A -	(C) <i>GPIP</i> . Nonsynonymous substitutions are indicated in boldface letters.

clearly supporting independent domestication of pigs in Europe and Asia.

We observed five nucleotide substitutions in a segment of 784 bp ( $K = 6.4 \pm 2.8 \times 10^{-3}$  subs/site) between the *GPIP\*3* and *\*4* alleles, predominantly found in Chinese Meishan and European domestic pigs, respectively. The nucleotide substitution rate for pseudogenes is  $\sim 4 \times 10^{-9}$  substitutions/site/year (Li 1997), suggesting that these two alleles diverged  $\sim$ 800,000 YBP. An estimated time since divergence of 2.7 million years was obtained for the two TYR alleles (differing by four synonymous substitutions,  $K_{\rm S} = 1.91 \pm 1.1 \times 10^{-2}$  subs/ site) using an average rate of synonymous substitutions for mammalian protein coding genes  $(3.51 \times 10^{-9} \text{ subs}/$ site/year; Li 1997). It should be noted that the confidence intervals for these two estimates are large and that they provide maximum estimates since part of the sequence divergence between the alleles may have occurred prior to the divergence of European and Asian populations.

The presence of allelic variants with a fairly large sequence divergence ( $\sim 0.5\%$ ) and the absence of intermediate forms at all three nuclear genes suggest that these alleles may have a European and Asian origin. If so, the putative Asian alleles found among European wild boars must reflect some gene flow from domestic to wild pigs in Europe. It would not be surprising if such gene flow occurred during the last 200 years. Alter-

natively, the alleles represent ancestral polymorphism									
showing significant allele frequency differences be-									
tween European and Asian populations.									

#### DISCUSSION

This study has important implications for pig genetics and the history of domestication. We report clear evidence for independent domestication of both European and Asian subspecies of the wild boar. We estimated the time since divergence of the ancestors for European and Chinese Meishan domestic pigs at  $\sim$ 500,000 YBP on the basis of interpopulational distances for mtDNA *cytB* sequences, which we think is the most reliable and conservative estimate we have obtained. It is clear that this is still a very rough estimate for several reasons. First, there is a fairly large sampling error since we have studied only  $\sim 10\%$  of the total mtDNA genome and a few nuclear genes. However, since the submission of this article, we have sequenced the entire mtDNA and can confirm the existence of the three major clades EI, EII, and A (Figure 2) and the large distance between European clade E1 and Asian clade A ( $K = 1.2 \pm 0.09\%$ ; J. H. M. Kijas and L. Andersson, unpublished results). Second, the molecular clock rate shows a considerable variation among lineages, possibly due to differences in efficiency in DNA repair, generation time, and/or metabolic rate (Li 1997). However, we can safely con-

**TABLE 2** 

		$T_{i}$	YR		GPIP						
Population	n	*1	*2	*1	*2	*3	*4				
Asia											
Wild boar, Japan	7	0.93	0.07	0.71	0.00	0.29	0.00				
Meishan	7	0.93	0.07	0.00	0.00	0.93	0.07				
Europe											
Wild boar, Italy	20	0.05	0.95	0.00	0.00	0.04	0.96				
Wild boar, Poland	13	0.45	0.55	0.00	0.00	0.12	0.88				
Large White	20	0.35	0.65	0.00	0.00	0.27	0.73				
Landrace	13	0.45	0.55	0.00	0.00	0.27	0.73				
Hampshire	19	0.00	1.00	0.00	0.00	0.24	0.76				
Duroc	10	0.23	0.77	0.00	0.00	0.10	0.90				
Other areas											
Wild boar, Israel	3	0.50	0.50	0.00	0.00	0.00	1.00				
Domestic pig, Cook Island	1	0.00	1.00	0.00	1.00	0.00	0.00				

Allele frequencies at the TYR and GPIP loci among domestic pig and wild boar populations

clude that the time since divergence must predate the domestication of pigs  $\sim$ 9000 YBP. Our results agree with Darwin's view (Darwin 1868) that there is a considerable genetic distance between European and Asian domestic pigs but they disagree with the hypothesis that Asian pigs originate from a wild boar species (*S. indicus*) unknown from the wild. Our data show that European domestic pigs and Chinese Meishan pigs are closely related to existing subspecies of the Eurasian wild boar (S. scrofa). The marked morphological divergence between Asian and European domestic pigs noted by Darwin may be due to the long history of selective breeding. Our results are in sharp contrast to the estimate of  $\sim$ 2000 years of divergence between Chinese Meishan pigs and European domestic pigs based on allele frequencies at six linked microsatellite loci (Paszek et al. 1998). However, this latter approach is invalid if introgression between the two forms has occurred.

The molecular data presented here, clearly indicating an introgression of Asian domestic pigs into European breeds, are consistent with historical written records (Darwin 1868; Jones 1998). A hybrid origin of the European Large White breed was suggested previously on the basis of restriction site mapping of mtDNA (Watanabe et al. 1986). Our results indicate a surprisingly high proportion of Asian mtDNA in this breed. However, this may not truly reflect the degree of introgression because the frequency of mtDNA types may have changed by genetic drift and/or selection during the  $\sim$ 200 years that have passed since the major introgression occurred. The hybrid origin of some major breeds may in fact be a unique advantage for the genetic dissection of complex traits of agricultural significance. More or less intense pig breeding has been practiced for thousands of years in both Asia and Europe. Introgression of Asian pigs into certain European breeds occurred  $\sim$ 200 years ago and was followed by intense phenotypic selection. Thus, we expect Asian alleles with favorable phenotypic effects to be fixed or to occur at a high frequency in European pig breeds with a hybrid origin. This is well illustrated by the MC1R allele for dominant black color (with a presumed Asian origin) fixed in the European Large Black breed (Kijas et al. 1998). The large number of generations that has passed since major crossbreeding occurred (+100 generations) implies that linkage disequilibrium occurs only within very narrow regions of genomic DNA. Thus, these breeds with a hybrid origin may be considered as very advanced intercross lines suitable for high-resolution quantitative trait loci mapping (Darvasi and Soller 1995; Darvasi 1998). The possibility of using large numbers of single nucleotide polymorphisms provides the tool for detecting islands of linkage disequilibrium created by introgression and phenotypic selection. The hybrid origin of the Large White breed must also be taken into account when evaluating the results of pig gene mapping experiments involving crosses between

this breed and European wild boars (Andersson *et al.* 1994) or Chinese Meishan pigs (Rohrer and Keele 1998).

The data presented on the domestic pig add to an emerging picture of independent domestication of distinct populations of the wild ancestors of major domestic species. In cattle, a considerable genetic divergence between European/African (*Bos taurus taurus*) and Asian (Bos taurus indicus) cattle has been documented on the basis of both mtDNA and nuclear DNA sequences (Loftus et al. 1994; MacHugh et al. 1997). The presence of distinct European and Asian mtDNA clades in domestic sheep has also been reported (Hiendleder et al. 1998). The considerable genetic diversity detected in these species has been interpreted to largely reflect genetic divergence prior to domestication. In contrast, ancient domestication (>100,000 YBP) was invoked to explain the pattern of genetic diversity in the control region of mtDNA among dogs (Vilá et al. 1997). The average genetic distance between the European and Asian clades of mtDNA *cytB* sequences was estimated at 1.5%, which is about three times higher than the 0.5% observed between the most divergent human mtDNA haplotypes (Horai et al. 1995). The sequence divergence between the alleles at the three nuclear genes included in this study is also several times higher than that generally observed in humans (Chakravarti 1999). These differences in pattern of genetic diversity are consistent with a strikingly different population history of humans and domestic animals. Humans appear to have expanded enormously from a rather small population present 100,000 YBP (Harpending et al. 1998). At that time, the wild boar was abundant and widespread in Eurasia. The independent domestication of European and Asian wild boars followed by introgression provided a broad genetic basis for the domestic pig.

The molecular markers described here give an opportunity to further explore the origin of domestic pigs. It is clear that a more thorough sampling of domestic breeds, particularly from Asia, and of wild boar populations is needed to provide a more comprehensive picture of the domestication of pigs. It would be of particular interest to include wild boar populations with a karyotype of 2n = 36 in contrast to 2n = 38 in other wild boar populations as well as in all domestic pigs. It will now be possible to trace the origin of different breeds of pigs as well as feral pigs present in North and South America, Australia, and on Pacific islands. This study gives a first glimpse of what this type of study may reveal. The single pig tested from Cook Island in Polynesia appears to belong to a hybrid population since the mtDNA and TYR sequence was identical to those predominantly found in European pigs whereas the GPIP allele was almost identical to the allele found in Chinese Meishan pigs. The result is not unexpected as the ancestral Polynesians brought domestic pigs with them (Diamond 1997, p. 60), presumably of Asian origin, and European seafarers used domestic pigs as onvoyage food supply during expeditions around the world. Such pigs sometimes escaped or were released and founded feral populations. Moreover, it may be possible to investigate the early history of the domestic pig by analyzing mtDNA from the rich collections of pig bones from Neolithic to medieval times. This may also shed light on the early migration of agricultural tribes in Eurasia.

We thank C. Vilá for valuable discussions, U. Gustafsson for excellent technical assistance, and R. Eiger, E. Geffen, T. Hori, R. Giannatelli, J. Kuryl, E. Randi, L. Varga, and the PiGMaP consortium for providing tissue or DNA samples. The Swedish Research Council for Forestry and Agriculture supported the study.

#### LITERATURE CITED

- Andersson, L., C. S. Haley, H. Ellegren, S. A. Knott, M. Johansson et al., 1994 Genetic mapping of quantitative trait loci for growth and fatness in pigs. Science 263: 1771–1774.
- Archibald, A. L., C. S. Haley, J. F. Brown, S. Couperwhite, H. A. McQueen *et al.*, 1995 The PiGMaP consortium linkage map of the pig (*Sus scrofa*). Mamm. Genome **6**: 157–175.
- Bökönyi, S., 1974 History of Domestic Mammals in Central and Eastern Europe. Akademiai Kiado, Budapest.
- Brown, W. M., M. George, Jr., and A. C. Wilson, 1979 Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA 76: 1967–1971.
- Chakravarti, A., 1999 Population genetics—making sense out of sequence. Nat. Genet. 21: 56–60.
- Darvasi, A., 1998 Experimental strategies for the genetic dissection of complex traits in animal models. Nat. Genet. 18: 19–24.
- Darvasi, A., and M. Soller, 1995 Advanced intercross lines, an experimental population for fine genetic-mapping. Genetics 141: 1199–1207.
- Darwin, C., 1868 The Variation of Animals and Plants under Domestication. John Murray, London.
- Diamond, J., 1997 Guns, Germs and Steel. Jonathan Cape, London.
- Harbitz, I., B. P. Chowdhary, S. Kran and W. Davies, 1993 Characterization of a porcine glucosephosphate isomerase-processed pseudogene at chromosome 1q1.6-1.7. Mamm. Genome 4: 589– 592.
- Harpending, H. C., M. A. Batzer, M. Gurven, L. B. Jorde, A. R. Rogers *et al.*, 1998 Genetic traces of ancient demography. Proc. Natl. Acad. Sci. USA **95**: 1961–1967.
- Hiendleder, S., H. Lewalski, R. Wassmuth and A. Janke, 1998 The complete mitochondrial DNA sequence of the domestic sheep (Ovis aries) and comparison with the other major ovine haplotype. J. Mol. Evol. **47**: 441–448.
- Horai, S., K. Hayasaka, R. Kondo, K. Tsugane and N. Takahata, 1995 Recent African origin of modern humans revealed by com-

plete sequences of hominoid mitochondrial DNAs. Proc. Natl. Acad. Sci. USA **92**: 532–536.

- Irwin, D. M., T. D. Kocher and A. C. Wilson, 1991 Evolution of the cytochrome b gene of mammals. J. Mol. Evol. 32: 128–144.
- Jones, G. F., 1998 Genetic aspects of domestication, common breeds and their origin, pp. 17–50 in *The Genetics of the Pig*, edited by A. Ruvinsky and M. F. Rothschild. CAB International, Oxon, UK.
- Kijas, J. M. H., R. Wales, A. Tornsten, P. Chardon, M. Moller *et al.*, 1998 Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. Genetics 150: 1177–1185.
- Kimura, M., 1980 A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111–120.
- Kumar, S., K. Tamura and M. Nei, 1993 MEGA: Molecular Evolutionary Genetics Analysis, Version 1.0. The Pennsylvania State University, University Park, PA.
- Li, W. H., 1997 Molecular Evolution. Sinauer Associates, Sunderland, MA.
- Loftus, R. T., D. E. MacHugh, D. G. Bradley, P. M. Sharp and P. Cunningham, 1994 Evidence for two independent domestications of cattle. Proc. Natl. Acad. Sci. USA 91: 2757–2761.
- MacHugh, D. E., M. D. Shriver, R. T. Loftus, P. Cunningham and D. G. Bradley, 1997 Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). Genetics **146**: 1071–1086.
- Markl und, S., R. Chaudary, L. Markl und, K. Sandberg and L. Andersson, 1995 Extensive mtDNA diversity in horses revealed by PCR-SSCP analyses. Anim. Genet. 26: 193–196.
- Nei, M., 1987 Molecular Evolutionary Genetics, equation 10.21, p. 276. Columbia University Press, New York.
- Okumura, N., N. Ishiguro, M. Nakano, K. Hirai, A. Matsui *et al.*, 1996 Geographic population structure and sequence divergence in the mitochondrial DNA control region of the Japanese wild boar (*Sus scrofa leucomystax*), with reference to those of domestic pigs. Biochem. Genet. **34**: 179–189.
- Paszek, A. A., G. H. Flickinger, L. Fontanesi, C. W. Beattie, G. A. Rohrer *et al.*, 1998 Evaluating evolutionary divergence with microsatellites. J. Mol. Evol. 46: 121–126.
- Rohrer, G. A., and J. W. Keele, 1998 Identification of quantitative trait loci affecting carcass composition in swine. I. Fat deposition traits. J. Anim. Sci. 76: 2247–2254.
- Ruvinsky, A., and M. F. Rothschild, 1998 Systematics and evolution of the pig, pp. 1–16 in *The Genetics of the Pig*, edited by A. Ruvinsky and M. F. Rothschild. CAB International, Oxon, UK. Swofford, D. L., 1998 *PAUP\* 4.0 Beta Version: Phylogenetic Analysis*
- Swofford, D. L., 1998 PAUP\* 4.0 Beta Version: Phylogenetic Analysis Using Parsimony (and Other Methods). Sinauer Associates, Sunderland, MA.
- Ursing, B. M., and U. Arnason, 1998 The complete mitochondrial DNA sequence of the pig (*Sus scrofa*). J. Mol. Evol. **47**: 302–306.
- Vilá, C., P. Savolainen, J. E. Maldonado, I. R. Amorim, J. E. Rice et al., 1997 Multiple and ancient origins of the domestic dog. Science 276: 1687–1689.
- Ward, P., B. L. Frazier, K. Dew-Jager and S. Pääbo, 1991 Extensive mitochondrial diversity within a single Amerindian tribe. Proc. Natl. Acad. Sci. USA 88: 8720–8724.
- Watanabe, T., Y. Hayashi, J. Kimura, Y. Yasuda, N. Saitou *et al.*, 1986 Pig mitochondrial DNA: polymorphism, restriction map orientation, and sequence data. Biochem. Genet. 24: 385–396.

Communicating editor: C. Haley