Inter-population and intra-population maternal lineage genetics of Lanyu pig (Sus scrofa) by analysis of mitochondrial cytochrome b and control region sequences

Y. N. Jiang, C. Y. Wu, C. Y. Huang, H. P. Chu, M. W. Ke, M. S. Kung, K. Y. Li, C. H. Wang, S. H. Li, Y. Wang and Y. T. Ju

J Anim Sci published online Mar 14, 2008;

Running head: Population genetic structure of Lanyu pigs

Inter-population and intra-population maternal lineage genetics of Lanyu pig (*Sus scrofa*) by analysis of mitochondrial cytochrome b and control region sequences[1]

Y. N. Jiang*, C. Y. Wu*, C. Y. Huang*, H. P. Chu[#], M. W. Ke*, M. S. Kung*, K. Y. Li*, C. H. Wang[ɔ], S. H. Li[†], Y. Wang[†], and Y. T. Ju*[2]

[*]Department of Animal Science and Technology, National Taiwan University, Taipei, Taiwan; [#]Taitung Animal Propagation Station, Livestock Research Institute, Taitung, Taiwan; [ɔ]Kaohsiung Animal Propagation Station, Livestock Research Institute, Kaohsiung, Taiwan; [†]Department of Life Science, National Taiwan Normal University, Taipei, Taiwan

[1]Acknowledgements: This work was supported by a grant from the National Science Council of Taiwan, Republic of China (NSC 94-2317-B-002-020 and NSC 96-2317-B-002-003) and Council of Agriculture of Taiwan, Republic of China [95AS-11.1.2-AD-U1 and 96AS-11.1.4-AD-U1]. We thank Dr. Harry Wilson for editing the manuscript.

[2]Correspondence: ytju@ntu.edu.tw

Published Online First on March 14, 2008 as doi:10.2527/jas.2007-0049

ABSTRACT

The Lanyu pig is an indigenous breed from the Lanyu Islet, to the southeast of Taiwan. Two herds of Lanyu pigs were introduced from the Lanyu Islet into Taiwan in 1975 and 1980. The current population of conserved Lanyu pigs consists of only 44 animals with unknown genetic lineage. The Lanyu pig possesses a distinct maternal genetic lineage remote from Asian and European pigs. The present study aimed to understand the phylogenetic relationship among conserved Lanyu, Asian, and European type pigs based on the *cytochrome b* coding gene, to ascertain the maternal lineage and genetic diversity within the conserved Lanyu pigs, and to address whether genetic introgression from exotic or Formosan wild pigs had occurred in the conserved Lanyu pigs. The entire mitochondrial genomes of both types of Lanyu pig comprise 2 ribosomal RNAs, 22 tRNAs, and 13 protein-coding genes. Only 2 haplotypes of mitochondrial DNA control region and *cytochrome b* were identified in the conserved Lanyu pig herds. When Maximum Likelihood trees were constructed, the Type I Lanyu mitochondrial genes formed a unique clade with a large pairwise distance of both *cytochrome b* and control region from Asian and European type breeds, Formosan wild pigs, and exotic breeds. Significant loss of genetic diversity of mtDNAs within the conserved Lanyu pigs was demonstrated by low haplotype and nucleotide diversities, supported by Fu and Li's D* neutrality test (1.44055; $P < 0.05$). The mtDNA control region sequences of extant pigs in the Lanyu Islet, however, showed high haplotype and nucleotide diversity, and clustered with exotic pigs. These results indicate no maternal lineage mtDNA gene introgression from Formosan wild pigs and introduced exotic pigs to conserved Type I Lanyu pigs, and a severe loss of heterozygosity of mtDNA in conserved Lanyu pigs. The remaining extant

pigs on the Lanyu Islet have been introgressed with exotic breeds. Strategies for future conservation of native Lanyu pigs are now even more urgent and important.

Key words: control region, *cytochrome b*, genetic diversity, Lanyu pig, mitochondrial DNA, phylogenetic relationship

INTRODUCTION

The Lanyu pig is an indigenous miniature pig breed with black coat color inhabiting the Lanyu Islet near Southeast Taiwan (Fig. 1). Two herds of Lanyu pigs were transferred to Taiwan for conservation purposes before 1980 (Chyr et al., 2001). Based on the polymorphism of mitochondrial DNA (mtDNA) control region, only 2 haplotypes of the control region have been identified in all of the conserved Lanyu pigs. The Lanyu pig with one haplotype (Type I) possesses a unique maternal genetic lineage distinct from Asian and European breeds, whereas Lanyu pigs of the other haplotype (Type II) are clustered in the major Asian pig clade (Wu et al., 2007), indicating that the original Lanyu pig possesses a unique genetic background.

The Formosan wild pig (*Sus scrofa taivanus*) is a native pig in Taiwan (Chao and Fang, 1988). Some traits of Formosan wild pigs show striking phenotypic differences from Lanyu pigs, but their body conformation and small erect ears are similar. The extent of genetic relationship between the Lanyu and the Formosan wild pig is unknown. In addition, the degree and pattern of introgression of genetic material from the introduced exotic breeds into Lanyu pigs is also currently unknown.

As the population of conserved Lanyu pigs stood at only 44 animals in 2006 (Wu et al., 2007), this study was undertaken to assist in the future genetic conservation and recovery of this unusual breed. The phylogenetic relationship among conserved Lanyu, Asian, and European pig breeds was determined by analysis of the polymorphism of their *cytochrome b* sequences. The genetic diversity within the conserved population, and the presence of any genetic affinity among conserved Lanyu pigs, Formosan wild pigs, and exotic breeds in Taiwan, and the pigs currently extant in the Lanyu Islet were investigated by the diversity of mtDNA. We were also interested whether any maternal mtDNA

genetic introgression had previously occurred between the conserved Lanyu pig and the Formosan wild pig.

MATERIALS AND METHODS

Sample Collection and Preparation of Mitochondrial DNA

Conserved herd Lanyu pig blood samples were collected from all 39 pigs from the Taitung Animal Propagation Station (TAPS) and all 5 pigs from the National Taiwan University (NTU) teaching farm. Blood samples of exotic pigs were obtained from 4 Meishan, 12 Taoyuan, 10 Berkshire, 14 Landrace, 10 Yorkshire, and 5 Duroc pigs from reference stocks at the Taiwan Livestock Research Institute (TLRI). Twelve current Lanyu pig blood samples were collected separately from pigs of 6 different aboriginal tribes on the Lanyu Islet. Five Formosan wild pig blood samples were obtained from Taitung and Hualien counties (Figure 1). The mitochondrial DNAs were extracted and purified from platelet-rich plasma using Qiagen's QIAamp DNA mini kit (Qiagen, Valencia, CA). The quality of purified mtDNAs was analyzed via electrophoresis on a 1% agarose gel.

Primer Design and Amplification of mtDNA Fragments by Polymerase Chain Reaction

Entire sequences of the mtDNA control region were amplified by PCR in a MJ thermal cycler using the following primers: L1, 5'-CCAAGACTCAAGGAAGGAGA-3' (sequence of position 16542-16561 of pig mtDNA, GenBank accession number AF034253) and H1, 5'-GGCGCGGATACTTGCATGTG-3' (position 1290-1309). The

entire sequence of *cytochrome b* was amplified using the following primers: CbR1, 5'-GTCCTGCCCTGAGGACAA-3' (position 15735-15752); CbL1, 5'-GGTGCTG ATGGCGGAGTT-3' (position 16581-16564), and CbR2, 5'-CCAAGACTCAAGGAA GGAGA-3' (position 15363-15382); CbL2, 5'-GGCGCGGATACTTGCATGTG-3' (position 115-134). Thermal cycling was conducted in 50 µL volumes using the FastStart High Fidelity PCR system (Roche, Penzberg, Germany), each containing 1 ng of mtDNA, 10 mM Tris-HCl pH 8.3, 1.8 mM $MgCl_2$, 0.4 µM of each primer, 200 µM of each dNTP, and 2 units of FastStart polymerase. Thermal cycling parameters were as follows: 95°C for 5 min; 30 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 80 s; with a final extension at 72°C for 4 min. The resultant PCR products were then purified using the PCR-M™ clean up system (Viogene, Taipei). The complete control regions were sequenced in both directions using the following primers: L1; H1; L2, 5'-CCTATGTACGTCGTGCATTA-3' (position 160-179); L3, 5'-TACTTCAGGACC ATCTCACC-3' (position 434-453); H2, 5'-AGTGTAAGTTAGGCTTATTG-3' (position 963-982); and H3, 5'-TTGTGGTAGATTGGCGTAAA-3' (position 1072-1091). The *cytochrome b* fragments were bidirectionally sequenced using the following primers: CbR1, CbL1, CbR2, and CbL2. All sequences were determined using an Applied Biosystems 3730 DNA sequencer and analyzed with SeqEd software (Perkin-Elmer, Applied Biosystems). Full sequences of the control region and *cytochrome b* were generated by overlapping forward and reverse sequences with SeqEdit software (DNASTAR, Madison, WI; see Hein and Stϕvlbaek, 1996). The following reference mtDNA control region sequences were obtained from NCBI GenBank: 6 Yorkshire (AM040633-AM040638), 4 Meishan (I, AY230821; II, AY230827; III, D17739; IV,

AB041474), 6 Taoyuan (I, AM040641-AM040645; II, AM040646), 1 Hampshire (AY574046), 1 Berkshire (AM040639), 4 Landrace (AM040613-AM040616), and 6 Duroc (AM040623-AM040628) pigs, a Japanese wild pig (AB015085), a Ryukyu wild pig (AB015087), and an Italian wild pig (AB015094). All of the above sequences (except Meishan and Hampshire pigs, and Japanese, Ryukyu, and Italian wild pigs) come from pigs reared in isolation at the TLRI. The *cytochrome b* sequences were also obtained from NCBI GenBank, comprising 5 Ryukyu wild pigs (AB015071-AB015075), 6 Japanese wild pigs (AB015065-AB015070), Satsuma (AB015076), Erhualian (AF486861), Tongcheng (AF486862), Wannanhua (AF486873), Taoyuan (DQ534707), Meishan (AB015077), Diannan (AF486869), Large White (AF486874), Berkshire (AY574045), Landrace (AF034253), Duroc (AY337045), Hampshire (AY574046), Yucatan miniature (AB015081), and 2 Italian wild pigs (AB015082-AB015083) (Watanobe et al., 1999; Kim et al., 2002; Yang et al., 2003).

Analysis of Full Length mtDNA Genome

The Type I and II mtDNA control regions of Lanyu pigs were identified after their mtDNA was obtained. Twenty pairs of primers were designed according to the Landrace mitochondrial genomic sequence AF034253 and the annealing conditions of PCR were as listed in Table 1. The PCR was performed using Long PCR Enzyme Mix (Fermentas, Hanover, MD) in 50 μL volumes with the following parameters: 94°C for 5 min; 32 cycles of 94°C for 30 s, annealing for 30 s, and 68°C for 90 s; with a final extension at 68°C for 10 min. The resultant PCR products were then purified using the PCR-M™ clean up system (Viogene, Taipei), before direct sequencing was performed. All of the

amplified sequences were confirmed by bidirectional sequencing.

Data Analyses

In the analysis of pairwise distance of control regions, the tandem repeat motif 'CGTGCGTACA' with a variable number of repeats in individuals, and the Type I and II Lanyu-specific repeated motifs (ACACAAACC and TAAAACACTTA, respectively) in the mtDNA control region were excluded from the analysis (Wu et al., 2007). Sequence alignment of control region and *cytochrome b* was performed using MegAlign multiple alignment software (DNASTAR, Madison, WI; Hein and Stϕvlbaek, 1996). Haplotype and nucleotide diversities within the conserved Lanyu pigs, exotic pigs, and the extant pigs on the Lanyu Islet were obtained with DNA Sequence Polymorphism (DnaSP) software version 4.10.9 (Rozas et al., 2003). The PHYLIP program package (version 3.66) was used to obtain the maximum likelihood phylogenetic tree (Felsenstein, 2006). The TREE-PUZZLE (version 5.0) software, based on the quartet puzzling method, was used to analyze confidence intervals of the phylogeny and likelihood-mapping analyses (Schmidt et al., 2002). The significance of the difference among pig groups was tested using 10,000 permutations in the quartet puzzling algorithm (Strimmer and von Haeseler, 1996). For Maximum Likelihood (ML) analysis, Modeltest version 3.6 software (Posada and Crandall, 1998) was used to determine the best fit model of the data, including nucleotide composition, substitution matrix among nucleotides, and proportion of invariant sites. Nodal supports of the ML tree were evaluated by bootstrap resampling (1,000 replications) using the fast heuristic search algorithm implemented in software PAUP4.0β10 (Swofford, 2002). Fu and Li's D* neutrality test values were determined

with DnaSP software.

RESULTS AND DISCUSSION

Variation of mtDNA Control Regions and Cytochrome b Sequences in Conserved Lanyu Pigs

We had previously studied the phylogenetic relationship among conserved Lanyu pigs and 43 Asian and European pig breeds based on the pairwise distance of mtDNA control region (Wu et al., 2007). Only 2 control region haplotypes were identified in all conserved Lanyu pigs in Taiwan. The Lanyu pigs reared at TAPS have a mixture of Type I (28 head) and Type II (11 head) mtDNA haplotypes, while all pigs (5 head) at the NTU teaching farm had the Type I haplotype. Here, to determine whether the sequence polymorphism of *cytochrome b* (*Cytb*) coding gene matched that seen in the noncoding control region, the variation of *Cytb* sequences among the conserved Lanyu pigs was analyzed. The mtDNA *Cytb* sequences were obtained from all 44 Lanyu pigs kept at TAPS and NTU. Consistent with the control region analyses, all of the Lanyu pigs' mtDNA *Cytb* sequences from TAPS and the teaching farm of NTU could be categorized into 2 types (Figure 2). The phenotypic conformation of both types of Lanyu pigs are indistinguishable. All 5 of the Lanyu pigs at the NTU teaching farm had the same *Cytb* haplotype (Type I), while 2 haplotypes were found in the Lanyu herd at TAPS (Figure 2). In the mtDNA control region, 18 polymorphic sites (2 haplotypes), including 17 transitions and 1 transversion, were found in the conserved Lanyu pigs by Wu et al. (2007). An additional 10 transitions and 1 transversion in *Cytb* were found in the

same populations in the present analysis (Fig. 2; Suppl. Fig. S1). Based on the diversity of *Cytb* sequences, Fu and Li's D* test for neutrality was used to estimate the genetic variation within the population of conserved Lanyu pigs, and the test value was determined to be 1.44055 ($P < 0.05$), indicating significant deviation from neutrality. Among all 44 conserved Lanyu pigs, the low haplotype numbers and nucleotide diversities shown by DnaSP software analysis ($h = 0.384 \pm 0.066$; $\pi = 0.0037 \pm 0.00064$) indicate severe genetic drift. This confirms that there has indeed been a significant loss of heterozygosity of mtDNA haplotypes in the population of all conserved Lanyu pigs, probably a result of the small population size and uncontrolled mating in the past. Additionally, studies of gene frequency and genotypes of 19 microsatellite markers have also revealed that severe nuclear gene drift has occurred in the conserved Lanyu pig population (Chang et al. unpublished data).

Lanyu mtDNA Type I Cytb Is Genetically Remote from Asian and European Breeds

We had shown earlier that the Lanyu Type I mtDNA control region sequence was genetically remote from Asian and European breeds (Wu el al., 2007) and were interested whether the *Cytb* coding region was similar in distance to the noncoding control region, and whether gene introgression had occurred between Lanyu and Formosan wild pigs. The Asian and European pig breeds' *Cytb* sequences were obtained from NCBI database (see Materials and Methods for details). In addition, 5 Formosan wild and 44 conserved Lanyu pigs' *Cytb* sequences were determined. Fifty nucleotide substitutions were identified in *Cytb* sequences from the aforementioned pig breeds (Fig. 2). The nucleotide substitutions in positions 15422, 15425, 15623, and 16323 were specifically found in

Type I Lanyu *Cytb* sequences. A unique transversion in position 15425 was also identified in Lanyu Type I (thymine) compared to the consensus sequence (adenine). The pairwise genetic distance among the *Cytb* sequences was determined and a Maximum Likelihood phylogenetic tree was constructed. The most appropriate model for this data set was found to be HKY+*I*(-ln = 1935.7043; K = 5; AIC = 3881.4087). The maximum likelihood estimates of base frequencies were: A: 0.3174; C: 0.2887; G: 0.1292; and T: 0.2647. The estimated proportion of invariable sites was 0.7964. A transition/transversion ratio of 10.4402 was used to obtain the Maximum Likelihood tree.

Two major clades (Asian and European clades) were recognized in the Maximum Likelihood tree (Figure 3). The treelike topology and phylogenetic signal were obtained by the quartet puzzling method (quartet puzzling support value, 80.2%) supporting the branch assignments in this phylogenetic tree (Suppl. Fig. S2). In the Asian clade, Type I Lanyu and Ryukyu wild pig sequences formed 2 separate subclades distinct from the Asian major subclade. Sequences from Japanese and Formosan wild pigs and the Satsuma, Meishan, Taoyuan, Diannan, Wannahua, Tongcheng, Erhualian, Large White, and Berkshire breeds clustered with the Type II Lanyu haplotype in the major Asian subclade.

Characteristics of Complete Mitochondrial Genome of Different Lanyu Haplotypes

The Lanyu Type I control region and *Cytb* haplotypes are unique sequences. We next compared the whole mitochondrial genome sequences to find variations in whole mtDNA genomes of both Lanyu haplotypes. After excluding the tandem repeat motifs (5'-CGTGCGTACA), the length of the Type I Lanyu mitochondrial genome was 16,491

253 nucleotides (nt; accession number: EF375877), and of the Type II Lanyu genome was
254 16,498 nt (accession number: DQ972936). The total lengths of the 2 types of mtDNA
255 control regions differed because of the varying numbers of ACACAAACC and
256 TAAAACACTTA repeat motifs in their control regions (Wu et al., 2007). With the 5' end
257 of the control region assigned as the first nucleotide of Type II sequence, the L-strand
258 replication origin was located at positions 6233 to 6269. Both types of Lanyu
259 mitochondrial genome encoded 37 genes, including 2 ribosomal RNAs (12S and 16S
260 rRNA), 22 tRNAs, and 13 protein-coding genes, as listed in Table 2. One hundred and
261 twenty-four nucleotide substitutions (107 transitions; 17 transversions) were identified
262 between the Lanyu sequences. All protein-coding genes in the mitochondrial genome
263 used identical start and termination codons in both Lanyu sequences.

264 A triplicate ACACAAACC motif was specifically found in the mtDNA control region
265 of the Type I Lanyu sequence, whereas only 1 ACACAAACC motif was found in Type II
266 sequences and in exotic pigs. Another duplicate TAAAACACTTA motif in domain II of
267 the control region was specifically found in Type II Lanyu, whereas only 1
268 TAAAACACTTA motif was seen in Type I Lanyu and in exotic pigs (except Satsuma
269 breed; Wu et al., 2007). These 2 repeated motifs might be due to heteroplasmy, and result
270 in different lengths of full mitochondrial genomes in the 2 haplotypes of Lanyu pigs.

271 A typical animal mitochondrial genome encodes 36 to 37 genes, which are powerful
272 markers for inferring phylogenetic relationships (Saccone et al., 2000; Gerber et al.,
273 2001). The numbers of tRNA (22), ribosomal RNA (2), and protein-coding genes (13) in
274 both haplotypes of Lanyu mtDNA are identical to the numbers in domestic pigs (Ursing
275 and Arnason, 1998; Lin et al., 1999). All of the initiation codons (9 ATG, 3 ATA, and 1

GTG) of the 13 protein-coding genes were ATG, except the codons for *NADH dehydrogenase subunit 2* (*NADH2*), *NADH3*, and *NADH5*, which were ATA, and for *NADH4L*, which was the rare GTG. The ATA initiation codon for the *NADH2* gene was found in both types of Lanyu mtDNA, which differs from the rare ATT initiation codon in Landrace and Swedish domestic pigs (Ursing and Arnason, 1998; Lin et al., 1999). Some other slight differences noted were: the termination codons of 6 and 2 of 13 protein-coding genes used TAA and TAG, respectively, except those for *NADH3, NADH4, cytochrome c oxidase subunit II* (*COII*), *COIII*, and *Cytb*, which were TAT, TAC, TCA, TAC, and AGA, respectively (Table 2); 4 genes (*COII*, *COIII*, *NADH3*, and *NADH4*) terminated with an incomplete TNN stop codon, where NN was the 5' terminus of the adjacent tRNA gene. The incomplete stop codon (TNN) forms a stop codon by post-transcriptional polyadenylation (Anderson et al., 1981; Ojala et al., 1981; Wolstenholme, 1992).

The vertebrate control region is subdivided into 3 domains. The central domain of the control region, containing the replication origin of the heavy strand, is relatively well conserved. The 2 regions (domain I and II) flanking the central domain are hyper-variable in base substitution, insertion, and deletion (Saunders and Edwards, 2000). The complete domestic pig mtDNA has been sequenced and the control region is located between genes encoding tRNA-Pro and tRNA-Phe, containing approximately 1,245 nucleotides (Ursing and Arnason, 1998; Lin et al., 1999).

Phylogenetic Comparison of Exotic, Formosan Wild, and Conserved Lanyu Pigs

Many exotic pig breeds (including Taoyuan, Meishan, Berkshire, Yorkshire,

300 Landrace, and Duroc breeds) were introduced into Taiwan to improve the production
301 performance of local pigs, by colonizers from 120 to 50 yr ago and by the Taiwan
302 government in more recent times (Chyr et al., 2001). To explore whether exotic breeds'
303 and Formosan wild pigs' genes had introgressed into the conserved population of Lanyu
304 pigs, the pairwise distance and Maximum Likelihood methods were used to investigate
305 the phylogenetic relationship among the conserved Lanyu pigs, the exotic breeds, and
306 Formosan wild pigs. Some of the mtDNA control region sequences of exotic breeds were
307 obtained from the NCBI website, and mtDNA samples of further individuals including 12
308 Taoyuan, 4 Meishan, 10 Berkshire, 5 Duroc, 14 Landrace, and 10 Yorkshire breeds were
309 obtained from TLRI (Suppl. Fig. S3). The variable sites of the mtDNA control region of
310 conserved Lanyu, Formosan wild, and exotic pigs are listed in the supplemental data
311 (Suppl. Fig. S3). Unique nucleotide substitutions, including transitions at nucleotide
312 positions 302, 391, 535, 542, and 657, and a transversion at position 871 (thymine in Type
313 I Lanyu and adenosine in consensus sequence), were found in the control region of Type I
314 Lanyu mtDNA, but not other sequences used in this study. Pairwise distance analysis of
315 mtDNA haplotypes was again performed using DnaSP. A phylogenetic Maximum
316 Likelihood tree was constructed using the PHYLIP program package (Figure 4), which
317 showed that all pig sequences clustered into 3 major clades. Sequences from many
318 European pig breeds, including Hampshire, Landrace, Duroc, and Italian wild pig were
319 categorized as one major clade (referred to here as the European pig clade). Most other
320 sequences, including those from Formosan wild, Japanese wild, Ryukyu wild, Taoyuan,
321 Meishan, Yorkshire (Large White), Berkshire breeds, and the Type II Lanyu sequence,
322 were assigned to another major clade (referred to here as the Asian pig clade). The Type I

Lanyu sequence clustered as a unique clade distinct from the 2 major clades mentioned above, indicating that the maternal lineage of pigs containing the Type I Lanyu sequence had never crossbred with the Formosan wild and the exotic breeds. The pairwise distance shows that the 2 Lanyu sequences are very different from each other in their maternal lineages.

The treelike topology and phylogenetic signal (quartet puzzling support value, 86.2%) obtained by the quartet puzzling method supported the branches in this phylogenetic tree (Suppl. Fig. S4). In the maximum likelihood analysis, the most appropriate model for this data set was found to be TIM+*I*(-ln = 2003.8291; K = 7; AIC = 4021.6582). The maximum likelihood estimates of base frequencies were: A: 0.3639; C: 0.2662; G: 0.1230; and T: 0.2470. Estimated symmetrical substitution rates among these nucleotides were 1.0000 for A/C, 106.7043 for A/G, 47.4006 for T/C, 5.1202 for A/T, 5.1202 for C/G, and 1.0000 for G/T. The estimated proportion of invariable sites was 0.9109. A transition/transversion ratio of 12.5899 was used to obtain the Maximum Likelihood tree.

Distinct Genetic Lineage of Lanyu, Taoyuan, and Formosan Wild Pigs

We had previously shown that the Taoyuan (accession no. AM040645, AM040646) and Lanyu pigs possess distinct control region haplotypes (Wu et al., 2007). Here, we obtained an identical result after we increased the number of individual Taoyuan pigs in our analysis, indicating no mtDNA gene introgression between Taoyuan and Lanyu pigs (Figure 4). In the present study, we found a remote pairwise distance (0.01882 ± 0.00755 and 0.00936 ± 0.00102) of both control region and *Cytb* coding sequences between Type I Lanyu and Formosan wild pigs. Based on pairwise distance of *Cytb*, the Formosan wild

346 pigs were clustered together with Japanese wild pigs with 0.00361± 0.00120 pairwise
347 distance, while the Type II Lanyu sequence clustered with the major Asian breeds
348 subclade with 0.00328 ± 0.00072 pairwise distance. This result indicates no mtDNA
349 introgression between the Formosan wild pig and Lanyu pigs. The Diannan breed has a
350 similar phenotype to the Lanyu pig, but its *Cytb* sequence was identical to the *Cytb*
351 sequence in Formosan and Japanese wild pigs, suggesting that the mtDNA of East Asian
352 wild pigs might have introgressed into the Diannan breed.
353

354 ***Gene Introgression from Exotic Pig Breeds into the Pigs Extant on the Lanyu***
355 ***Islet***

356 The Lanyu Islet was isolated during the aboriginal culture protection period by the
357 Taiwan government. When it was opened to tourism and unrestricted travel after 1960, as
358 most of the Lanyu Islet pigs were bred in free-range piggeries, there was significantly
359 increased opportunity for the introduction of and crossbreeding with exotic pig breeds
360 from Taiwan. To understand the current diversity of mtDNA in pigs distributed
361 throughout the Lanyu Islet, we obtained mtDNAs from 12 individual Lanyu pigs reared
362 by 6 tribes on the Lanyu Islet during February 2005. Their mtDNA control region
363 sequences were subjected to phylogenetic analysis and compared to the sequences of
364 Lanyu pigs conserved in Taiwan, Formosan wild pigs, and exotic pigs in Taiwan. A
365 Maximum Likelihood tree was constructed and sequences from the modern pigs from the
366 Lanyu Islet were clustered into 3 groups (Figure 4). Four extant pigs on the Lanyu Islet
367 had control region sequences identical to the Type I Lanyu haplotype. Two pigs had
368 sequences that clustered together with Meishan I and II, Yorkshire, Japanese wild pig,
369 Formosan wild pigs, and Type II Lanyu sequences, but these 2 sequences had lost 1 of the

370 repeated TAAAACACTTA motifs present in duplicate in the Type II Lanyu haplotype.
371 The remaining 6 extant Lanyu pigs had mtDNA sequences that were grouped together
372 with Taoyuan, Berkshire, and Meishan III, IV sequences (Figure 4). This confirms that
373 most extant Lanyu pigs on the Lanyu Islet have hybridized with Taoyuan, Berkshire,
374 Meishan, or Yorkshire exotic pig breeds.

375 Most pigs extant on the Lanyu Islet have dark-pigmented coat color, and some of
376 them present phenotypes of exotic domestic breeds that are now or were previously found
377 in Taiwan, such as the Landrace, Yorkshire, Hampshire, Duroc, Taoyuan, Meishan, and
378 Berkshire breeds. The present study supports the hypothesis of recent introgression of
379 Meishan, Taoyuan, Berkshire, and Yorkshire genes into the extant pigs on the Lanyu Islet.
380 Although 4 extant pigs on the Lanyu Islet had identical mtDNA control region sequences
381 to Type I Lanyu, we observed that those pigs possessed physical characteristics typical of
382 Hampshire, Landrace, and Duroc phenotypes, suggesting that the introgression of exotic
383 pig genes is more serious than could be detected by a simple comparison of the variation
384 of maternal-linked mtDNAs. No extant pigs on the Lanyu Islet were found to possess the
385 duplicate TAAAACACTTA motif in their control regions, suggesting that the Type II
386 Lanyu sequence might be becoming extinct on the Lanyu Islet, its original habitat. One
387 pig from the Lanyu Islet actually possessed 4 ACACAAACC repeat motifs, indicating
388 either a recent mutation or that 4 ACACAAACC repeat motifs might have existed in
389 Lanyu pigs previously. These results demonstrated significant genetic introgression from
390 exotic pig breeds into the pigs extant on the Lanyu Islet, and that gene drift is currently
391 occurring on the Lanyu Islet, resulting in the loss of the Type II Lanyu sequence.
392

393 ***Evolution of Lanyu Pigs***

394 Based on the mtDNA sequences and morphometric data from museum specimens,
395 and tissue samples, Lucchini et al. (2005) presented a possible scenario for pig speciation
396 in South-east Asian (SEA): The SEA ancestral pig species (genus *Sus*) might have
397 crossed from Sundaland to the Philippines during the Pliocence (5.3-1.8 Mya). Larson et
398 al. (2005) constructed a consensus tree based on mtDNA obtained from 686 wild and
399 domestic pig specimens from museums worldwide, and showed that basal lineages
400 (origin) of *Sus scrofa* occurred on the western island Southeast Asia (ISEA) and dispersed
401 into the Indian subcontinent, then dispersed northward into the Asian continent, followed
402 by subsequent westward radiations into mainland Asia, and a final, progressive dispersal
403 across Eurasia into Western Europe. More than 1 domestication event [2 in China (Gansu
404 and Hunan provinces), 1 in India, 1 in Burma and Thailand, and another in Cape York of
405 Northern Australia], followed by rapid radiative expansion in Asia, was identified by
406 median-joining network analysis of Asian domestic and wild pigs mtDNA haplotypes
407 (Larson et al., 2005). Later on, Larson et al. (2007) investigated human-mediated *Sus*
408 dispersal in ISEA based on the polymorphism of 781 mtDNA sequences from modern
409 and ancient *Sus* specimens. They concluded that the endemic pigs in Taiwan were a
410 genetic link with mainland East Asian pigs, Micronesian, and the Philippine pigs, but not
411 with the pigs in Brunei, Sumatra, Oceania, and Polynesia.

412 Based on variation of the control region, the pairwise distances of the Type I Lanyu
413 haplotype versus Asian and European pigs were 0.01726 ± 0.00275 and 0.02216 ±
414 0.00889, respectively (Wu et al., 2007). The Type I sequence formed a unique clade
415 distinct from the major Asian clade and the European clade in the constructed Maximum
416 Likelihood tree. In addition, the pairwise distances between Type II Lanyu control region

417 versus Type I Lanyu control region (0.01744 ± 0.00125), Type II Lanyu sequence versus
418 Asian clade (0.00471 ± 0.00054) and Type II Lanyu versus European clade (0.01941 ±
419 0.00356), reveal apparent divergence of Type I and Type II Lanyu mtDNA control region
420 sequences (Wu et al., 2007). In the present study, the pairwise distances of both
421 haplotypes of Lanyu *Cytb* sequences versus Asian and European breeds were consistent
422 with the pairwise distances of the control region variants. Looking at our results together
423 with those of others suggests that the Type I and II mtDNA haplotypes had quite different
424 origins, as evidenced by the large calculated genetic distance between the 2 types, and the
425 fact that the 2 types specifically possess different numbers of the repeated
426 ACACAAACC and TAAAACACTTA motifs in their control region. The formation of
427 unique Type I haplotype mtDNA on Lanyu Islet happened earlier than the formations of
428 Type II haplotype and Formosan wild pigs' haplotypes. The ancestor of Type I Lanyu pigs
429 might have crossed through Sundaland into Taiwan and the Lanyu Islet before the last
430 glacial maximum (Meijaard, 2003) and evolved in isolation in Taiwan and the Lanyu Islet
431 after the last glacial period ended. To address the origin of the Type II haplotype, the Type
432 II mtDNA haplotype was aligned with 172 *Sus* mtDNA control region sequences
433 published by Larson's group (Larson et al., 2005; 2007). The Type II DNA sequence was
434 identical to the Sarawak (Malaya, DQ779294) specimen and differed by 1 base pair from
435 the Guizhou Xiang (China, AY486118) and Guam D (Northern Mariana Island,
436 AY884677), and differed by 2 base pairs from the Andaman (India, AY884705), Large
437 Black (Australia, AY463075), Kune Kune (New Zealand, AY463076), and Satuma
438 (Japan, AB015091) breeds. Our result revealed the Type II Lanyu pig in Taiwan shared
439 genetic lineage with pigs distributed in ISEA and Oceania, which was missed in the

440 sampling by Larson's group (Larson et al., 2005; 2007). The short pairwise distance
441 between Type II and Asian breeds mtDNA (0.00471 ± 0.00054 in control region and
442 0.00328 ± 0.00072 in *Cytb*) shows that the Type II Lanyu mtDNA is genetically almost
443 identical to that of Asian breeds, suggesting that formation of Type II mtDNA might have
444 occurred in recent times. These results promote 3 hypotheses about the formation of the
445 Type II mtDNA of Lanyu pig. The first hypothesis is that the Type II mtDNA originated
446 in ISEA and spread northward into the main Asian continent and Japan, and then
447 eastward into Australia and New Zealand through Taiwan and the Lanyu Islet, mediated
448 by human migrations. The second hypothesis is that the Type II Lanyu sequence
449 originated during domestication in mainland Asia and then crossed through Taiwan and
450 the Lanyu Islet and spread into Japan, ISEA, and Oceania. The third hypothesis is that the
451 Type II Lanyu might have originated in Taiwan and Lanyu Islet and then spread into the
452 Asian mainland, Japan, ISEA, and Oceania. To evaluate the actual origin of the 2 mtDNA
453 haplotypes of Lanyu pigs in Asia, a further combination of genetic, biogeographic, and
454 zooarcheological data is needed.

455 The conserved Lanyu pigs possess a distinct maternal lineage to Asian and European
456 type breeds, with no evidence of mtDNA gene introgression from exotic and Formosan
457 wild pigs during recent times. The results of this study further emphasize their unique
458 maternal genetic characterization and the importance of understanding their genetic
459 origin in assessing the trajectories of prehistoric human migration from mainland East
460 Asia into ISEA. A significant loss of mtDNA diversity in the conserved Lanyu pigs
461 may be due to small population size and exotic gene introgressions. The discovery
462 that the majority of the pigs extant on the Lanyu Islet are hybrids indicates that

463 more efforts are needed for the recovery of this native breed. For further
464 conservation or restoration of the Lanyu pig as a distinct breed, the nuclear phylogenetic
465 relationship of remaining Lanyu pigs, Asian pigs, and other breeds throughout the world
466 will require further analysis. Future population management will also require deeper
467 analysis of global nuclear genetic characteristics within the population of conserved
468 Lanyu pigs using microsatellite markers or coding genes.
469
470
471 **LITERATURE CITED**
472
473 Anderson, S., A. T. Bankier, B. G. Barrell, M. H. L. de Bruijn, A. R. Coulson, J. Drouin,
474 I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R.
475 Staden, and I. G. Young. 1981. Sequence and organization of the human
476 mitochondrial genome. Nature 290:457-464.
477 Chao, J. T., and K. Y. Fang. 1988. The preliminary study on the biology of Formosan wild
478 pig, *Sus scrofa taivanus*. Taiwan J. Forest Sci. 3:353-362.
479 Chyr, S. C., K. J. Lin, H. L. Chang, T. S. Yang, and H. T. Yen. 2001. Guideline in Animal
480 Science-Swine Production. Chinese Society for Animal Science Publisher. Taipei.
481 Felsenstein, J. 2006. PHYLIP version 3.66 executables for PowerMac. University of
482 Washington, Seattle.
483 Gerber, A. S., R. Loggins, S. Kumar, and T. E. Dowling. 2001. Does nonneutral evolution
484 shape observed patterns of DNA variation in animal mitochondrial genomes? Annu.
485 Rev. Genet. 35:539-566.

Hein, J., and J. Stϕvlbaek. 1996. Combined DNA and protein alignment. Methods Enzymol. 266:402-418.

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

Larson, G., K. Dobney, U. Albarella, M. Fang, E. Matisoo-Smith, J. Robins, S. Lowden, H. Finlayson, T. Brand, E. Willerslev, P. Rowley-Conwy, L. Andersson, and A. Cooper. 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science 307:1618-1621.

Larson, G., T. Cucchi, M. Fujita, E. Matisoo-Smith, J. Robins, A. Anderson, B. Rolett, M. Spriggs, G. Dolman, T.-H. Kim, N. T. D. Thuy, E. Randi, M. Doherty, R. A. Due, R. Bollt, T. Djubiantono, B. Griffin, M. Intoh, E. Keane, P. Kirch, K.-T. Li, M. Morwood, L. M. Pedrina, P. J. Piper, R. J. Rabett, P. Shooter, G. Van den Bergh, E. West, S. Wickler, J. Yuan, A. Cooper, and K. Dobney. 2007. Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. Proc. Natl. Acad. Sci. U. S. A. 104:4834-4839.

Lin, C. S., Y. L. Sun, C. Y. Liu, P. C. Yang, L. C. Chang, I. C. Cheng, S. J. T. Mao, and M. C. Huang. 1999. Complete nucleotide sequence of pig (*Sus scrofa*) mitochondrial genome and dating evolutionary divergence within Artiodactyla. Gene 236:107-114.

Lucchini, V., E. Meijaard, C. H. Diong, C. P. Groves, and E. Randi. 2005. New phylogenetic perspectives among species of South-east Asian wild pig (*Sus* spp.) based on mtDNA sequences and morphometric data. J. Zool. 266:25-35.

Meijaard, E. 2003. Mammals of South-East Asian islands and their Late Pleiostocene

509 environments. J. Biogeogr. 30:1245-1257.
510 Ojala, D., J. Montoya, and G. Attardi. 1981. tRNA punctuation model of RNA processing
511 in human mitochondria. Nature 290:470-474.
512 Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA
513 substitution. Bioinformatics 14:817-818.
514 Rozas, J., J. C. Sánchez-Delbarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA
515 polymorphism analyses by the coalescent and other methods. Bioinformatics
516 19:2496-2497.
517 Saccone, C., C. Gissi, C. Lanave, A. Larizza, G. Pesole, and A. Reyes. 2000. Evolution of
518 the mitochondrial genetic system: an overview. Gene 261:153-159.
519 Saunders, M. A., and S. V. Edwards. 2000. Dynamics and phylogenetic implications of
520 mtDNA control region sequences in New World Jays (Aves: Corvidae). J. Mol. Evol.
521 51:97-109.
522 Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler. 2002. TREE-PUZZLE:
523 maximum likelihood phylogenetic analysis using quartets and parallel computing.
524 Bioinformatics 18:502-504.
525 Strimmer, K., and A. von Haeseler. 1996. Quartet puzzling: A quartet maximum-
526 likelihood method for reconstructing tree topologies. Mol. Biol. Evol. 13:964-969.
527 Swofford, D. L. 2002. PAUP*. *Phylogenetic Analysis Using Parsimony* (*and other
528 Methods), version 4.0 beta10. Sunderland: Sinauer Associates, Inc. Sunderland, MA.
529 Ursing, B. M., and U. Arnason. 1998. The complete mitochondrial DNA sequence of the
530 pig (*Sus scrofa*). J. Mol. Evol. 47:302-306.
531 Watanobe, T., N. Okumura, N. Ishiguro, M. Nakano, A. Matsui, M. Sahara, and M.

532 Komatsu. 1999. Genetic relationship and distribution of the Japanese wild boar (*Sus*
533 *scrofa leucomystax*) and Ryukyu wild boar (*Sus scrofa riukiuanus*) analysed by
534 mitochondrial DNA. Mol. Ecol. 8:1509-1512.
535 Wolstenholme, D. R. 1992. Animal mitochondrial DNA: structure and evolution. Int. Rev.
536 Cytol. 141:173-216.
537 Wu, C. Y., Y. N. Jiang, H. P. Chu, S. H. Li, Y. Wang, Y. H. Li, Y. Chang, and Y. T. Ju. 2007.
538 Type I Lanyu pig has a maternal genetic lineage distinct from Asian and European
539 pigs. Anim. Genet. 38:499-505.
540 Yang, J., J. Wang, J. Kijas, B. Liu, H. Han, M. Yu, H. Yang, S. Zhao, and K. Li. 2003.
541 Genetic diversity present within the near-complete mtDNA genome of 17 breeds of
542 indigenous Chinese pigs. J. Hered. 94:381-385.

TABLES

Table 1. PCR primers and annealing conditions

Primer pair	Locations in AF034253[1,2]	Annealing temperature (°C)
1F/1B	432-451/1398-1381	55
2F/2B	1313-1332/2401-2377	55
3F/3B	2199-2216/3278-3259	55
4F/4B	3150-3167/4112-4095	55
5F/5B	3983-4002/4846-4829	55
6F/6B	4754-4771/5849-5833	55
7F/7B	5625-5642/6893-6876	55
8F/8B	6579-6597/7736-7716	55
9F/9B	7629-7646/8685-8668	55
10F/10B	8584-8603/9568-9587	55
11F/11B	9437-9457/10418-10401	62
12F/12B	10002-10019/10879-10862	55
13F/13B	10515-10535/11584-11571	55
14F/14B	11530-11547/12592-12569	55
15F/15B	12459-12477/13381-13361	55
16F/16B	13057-13075/14156-14138	55
17F/17B	14089-14106/15057-15038	55
18F/18B	14930-14949/15843-15825	55
19F/19B	15735-15752/16581-16564	55
20F/20B	16485-16504/1398-1381	65

[1] The first nucleotide of the control region in H-strand is designated as nucleotide position 1.

[2] The nucleotide position numbers of each primer correspond to those in Landrace pig control region sequence AF034253.

553 **Table 2. Comparison of the mitochondrial genomes of Lanyu and Landrace pigs**

554

Feature	Position[1]			Strand[2]	Start codon			Stop codon		
	Lanyu		Landrace		Lanyu		Landrace	Lanyu		Landrace
	Type I	Type II			Type I	Type II		Type I	Type II	
D-loop	1-1062	1-1056	1-1175	H						
tRNA-Phe	1063-1132	1056-1125	1176-1245	H						
12S rRNA	1133-2090	1126-2087	1246-2205	H						
tRNA-Val	2089-2156	2087-2154	2206–2273	H						
16S rRNA	2157-3727	2155-3726	2274–3844	H						
tRNA-Leu (UUR)	3727-3801	3726-3800	3845–3919	H						
NADH1	3803-4762	3802-4758	3922–4878	H	ATG	ATG	ATG	TAG	TAG	TAG
tRNA-Ile	4761-4829	4757-4825	4877–4945	H						
tRNA-Gln	4828-4898	4825-4895	4943–5015	L						
tRNA-Met	4900-4969	4897-4966	5017–5086	H						
NADH2	4970-6013	4967-6010	5087–6130	H	ATA	ATA	ATT	TAG	TAG	TAG
tRNA-Trp	6012-6079	6009-6076	6129–6196	H						
tRNA-Ala	6086-6153	6083-6150	6203–6270	L						
tRNA-Asn	6155-6229	6152-6226	6272–6346	L						
O_L	6226-6272	6223-6269	6346–6382	L						
tRNA-Cys	6262-6327	6259-6324	6379–6444	L						
tRNA-Tyr	6327-6390	6324-6387	6444–6509	L						
COI	6394-7938	6391-7935	6511–8055	H	ATG	ATG	ATG	TAA	TAA	TAA
tRNA-Ser (UCN)	7942-8010	7939-8007	8059–8129	L						
tRNA-Asp	8018-8085	8015-8082	8135–8202	H						
COII	8086-8781	8083-8778	8203–8890	H	ATG	ATG	ATG	TCA	TCA	TNN
tRNA-Lys	8775-8840	8772-8837	8891–8957	H						
ATPase8	8842-9045	8839-9042	8959–9162	H	ATG	ATG	ATG	TAA	TAA	TAA
ATPase6	9003-9683	9000-9680	9120–9800	H	ATG	ATG	ATG	TAA	TAA	TAA
COIII	9683-10466	9680-10463	9800–10583	H	ATG	ATG	ATG	TAC	TAC	TNN
tRNA-Gly	10467-10535	10464-10532	10584–10652	H						
NADH3	10536-10892	10533-10889	10653–10998	H	ATA	ATA	ATA	TAT	TAT	TNN
tRNA-Arg	10883-10951	10880-10948	11000–11068	H						
NADH4L	10952-11248	10949-11245	11069–11365	H	GTG	GTG	GTG	TAA	TAA	TAA
NADH4	11242-12619	11239-12616	11359–12736	H	ATG	ATG	ATG	TAC	TAC	TNN
tRNA-His	12620-12688	12617-12685	12737–12805	H						
tRNA-Ser (AGY)	12689-12747	12686-12744	12806–12864	H						
tRNA-Leu (CUN)	12748-12817	12745-12814	12865–12934	H						
NADH5	12818-14641	12815-14638	12935–14755	H	ATA	ATA	ATA	TAA	TAA	TAA
NADH6	14623-15146	14627-15154	14739–15266	L	ATG	ATG	ATG	TAA	TAA	TAA
tRNA-Glu	15147-15215	15155-15223	15267–15335	L						
Cyt b	15220-16359	15227-16366	15342–16481	H	ATG	ATG	ATG	AGA	AGA	AGA
tRNA-Thr	16360-16427	16367-16434	16482–16549	H						
tRNA-Pro	16427-16491	16434-16498	16550–16613	L						

555 [1] The first nucleotide of the control region in H-strand is designated as nucleotide position 1.

556 [2] H and L indicate that the gene is transcribed from the H-strand or L-strand, respectively.

557 [3] TNN indicates the incomplete stop codon, where the NN is the 5' end of the adjacent tRNA nucleotide, which formed

558 a stop codon by post-transcriptional polyadenylation.

559 Abbreviations: NADH1-6 and NADH4L, subunits 1-6 and 4L of *nicotinamide dinucleotide dehydrogenase*; ATPase6

560 and 8, subunits 6 and 8 of *adenosine triphosphatase*; COI-COIII, *cytochrome c oxidase subunits I-III*; Cyt b,

561 *cytochrome b*.

562

563 **FIGURE LEGENDS**

564

565 **Figure 1: The Lanyu Islet pig breed.**

566 (A) A representative five month old male Lanyu pig. (B) Location of Taiwan and the

567 Lanyu Islet. (◆) and (*) represent the location of Hualien and Taitung county,

568 respectively.

569

570 **Figure 2: Variable sites of mtDNA *Cytb* sequences of Lanyu, Asian type, and**
571 **European type pig breeds.**

572 Sequence position number (given in the first row) follows those in the *Cytb* sequence of

573 the Landrace breed (accession number: AF034253). Only variable sites in the *Cytb* of

574 these animals, with sequence positions given above, are shown. Abbreviations in the

575 leftmost column indicate the geographical origins of these animals: T, Taiwan; J, Japan; E,

576 Europe; C, China. Nucleotides identical to the *Cytb* consensus sequence are denoted by a

577 dot (•). Nucleotides with filled blocks indicate the nucleotide specifically found in Type I

578 Lanyu haplotype.

579

580 **Figure 3: Phylogenetic tree of *Cytb* of conserved Lanyu, Asian and European pig**
581 **breeds.**

582 The phylogenetic tree was constructed on the basis of maximum likelihood distances of

583 polymorphism of *Cytb* sequences using the PHYLIP program package. The numeral on

584 the right side of the breed name represents the pig's identity number. Those branches with

585 highly significant and significant confidence are shown by bold and middle weight lines,

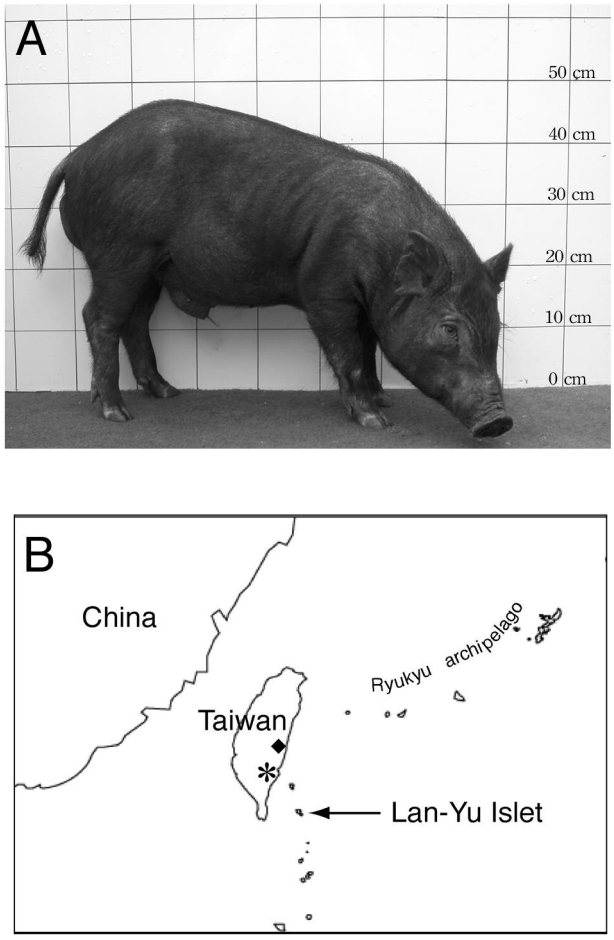
586 respectively. Numbers on the branches are bootstrap values based on bootstrap

587 resampling (1,000 replications).

588 **Figure 4: Phylogenetic tree of the mtDNA control region of conserved Lanyu pigs,**
589 **extant domestic Lanyu pigs, Formosan wild pigs, and exotic pigs in Taiwan.**
590 Imported exotic pig breeds include the Taoyuan, Meishan, Duroc, Landrace, Yorkshire,
591 Hampshire, and Berkshire breeds. The phylogenetic tree was constructed on maximum
592 likelihood distances of mtDNA control region polymorphism using the PHYLIP program
593 package. The numeral following the breed name represents the pig's identity number. The
594 number in parentheses indicates the sampling individuals used. (*) indicates the mtDNAs
595 from pigs reared in TLRI. Abbreviations: The Lanyu Islet indicates an extant pig from the
596 Lanyu Islet. Those branches with highly significant and significant confidence are shown
597 by bold and middle weight lines, respectively. Numbers on the branches are bootstrap
598 values based on bootstrap resampling (1,000 replications).

599

600 Figure 1.



601

602

Figure 2.

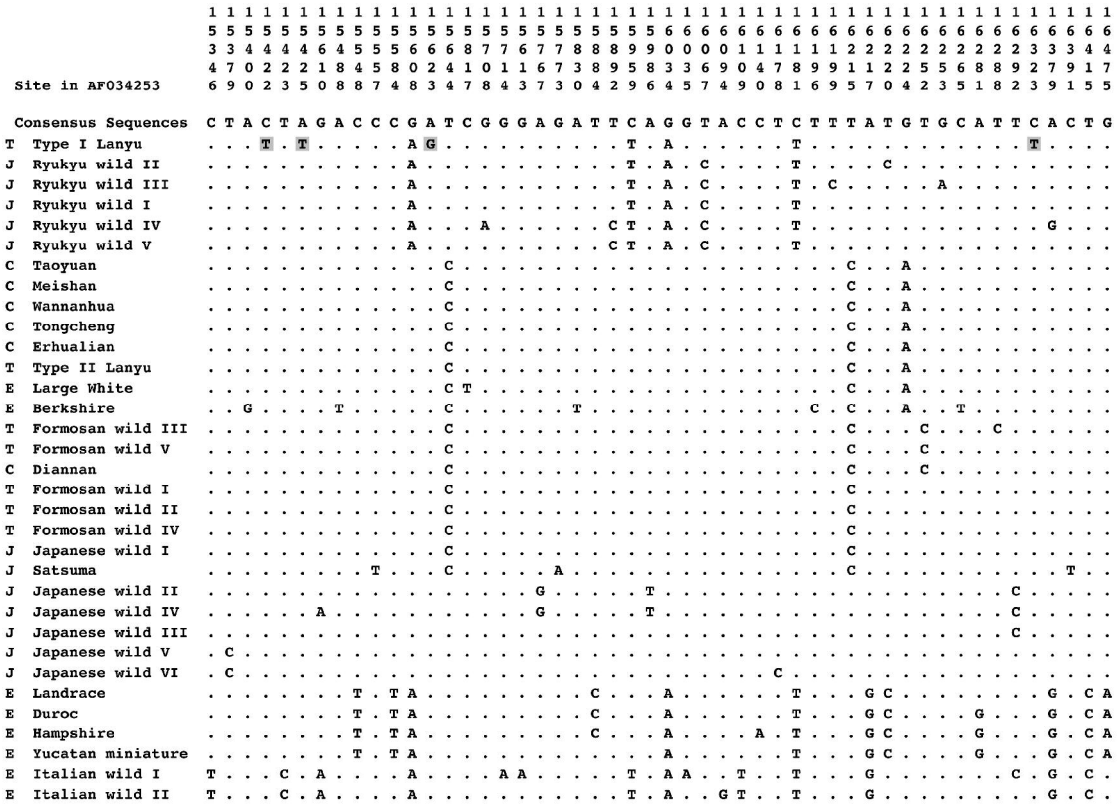


Figure 3.

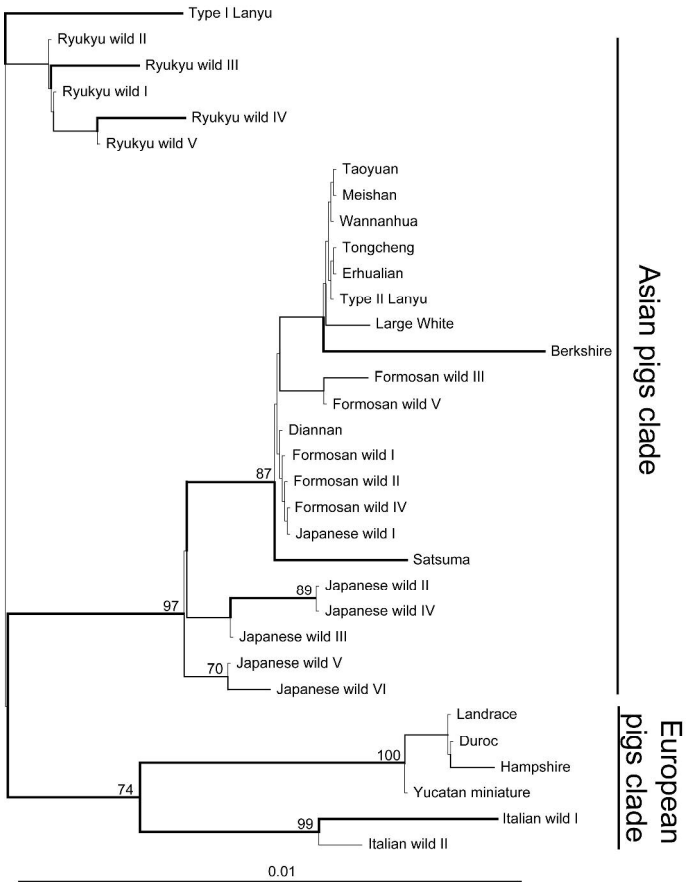
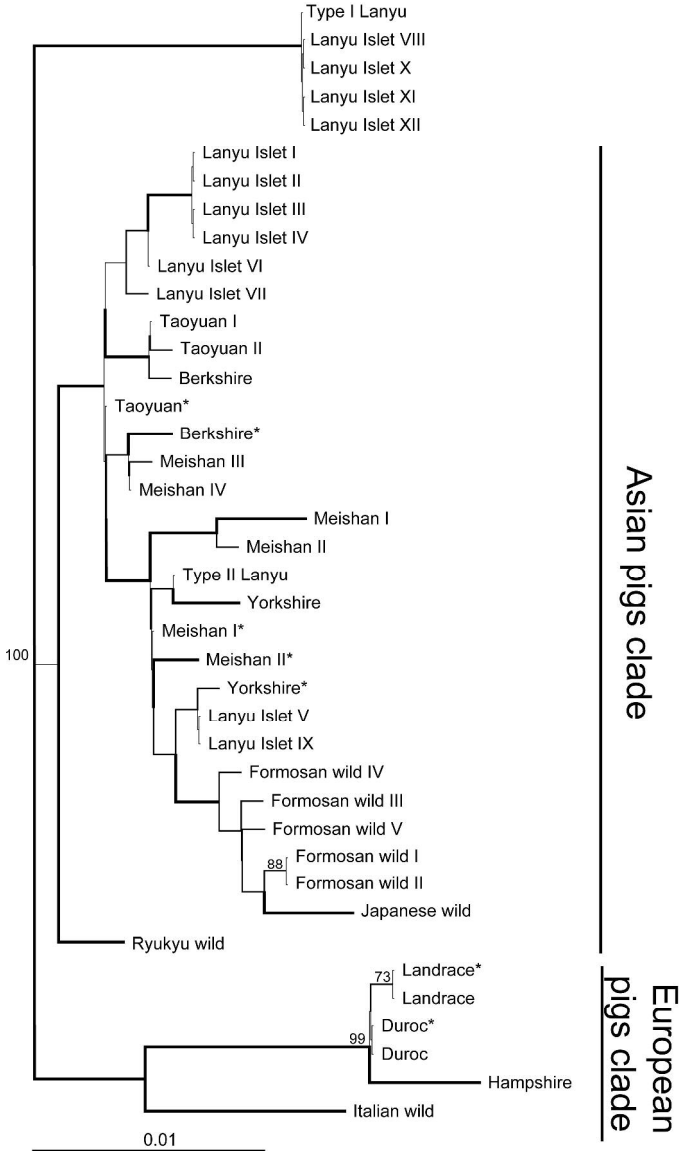


Figure 4.



Supplemental data

		Control region																	Cytochrome b										
	Haplotype															+														
		9	1	2	3	3	3	3	3	4	4	5	5	5	6	6	7	8	1	1	1	1	1	1	1	1	1	1	1
		0	8	7	0	0	0	2	9	0	5	3	4	7	5	9	0	7	0	5	5	5	5	5	5	6	6	6	6
			1	9	1	2	6	3	1	5	2	5	2	5	7	2	2	1	9	4	4	6	6	6	9	0	1	2	2	3
																			6	2	2	0	2	4	5	3	8	1	2	2
																				2	5	8	3	4	9	4	1	5	4	3
NTU (5)	I	A	T	T	T	G	C	C	T	T	C	A	G	A	T	A	A	T	A	T	T	A	G	T	T	A	T	T	G	T
TAPS (28)	I	A	T	T	T	G	C	C	T	T	C	A	G	A	T	A	A	T	A	T	T	A	G	T	T	A	T	T	G	T
TAPS (11)	II	G	C	C	C	A	T	T	C	C	T	G	A	G	C	G	G	A	G	C	A	G	A	C	C	G	C	C	A	C

Figure S1. The variation of mtDNA control region and *cytochrome b* nucleotide sequences in the two herds of conserved Lanyu pigs. Nucleotide variations in mtDNA control region and *cytochrome b* of conserved Lanyu pigs. Sequence breed codes and numbers are given in first row according to Landrace breed (accession number: AF034253). Abbreviations are as follows: NTU, Lanyu pigs from teaching farm of NTU. TAPS, Lanyu pigs from Taitung Animal Propagation Station. Only variable sites in mtDNA control region and *cytochrome b* of pigs, with sequence positions given above, are shown. The number inside the parentheses in left-hand column indicates the number of pigs obtained for mtDNA analysis. Plus signs (+) in position 702 represent substitutions occurring in repeat motifs. Nucleotides with filled blocks indicate the transversion nucleotides.

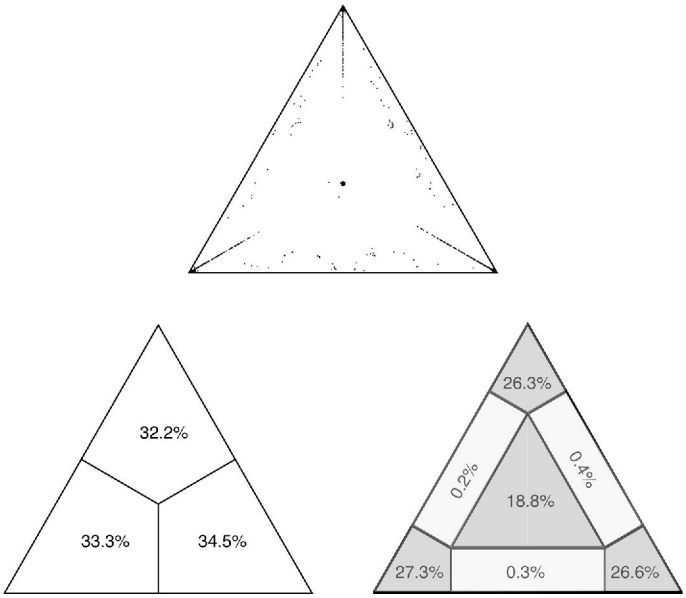


Figure S2. Quartet Puzzling analysis of Maximun Likelihood tree based on the pairwise distance of *cytochrome b*. The treelike topology and phylogenetic signal were obtained by the quartet puzzling method (quartet puzzling support value, 80.2%) supporting the branch assignments in this phylogenetic tree.

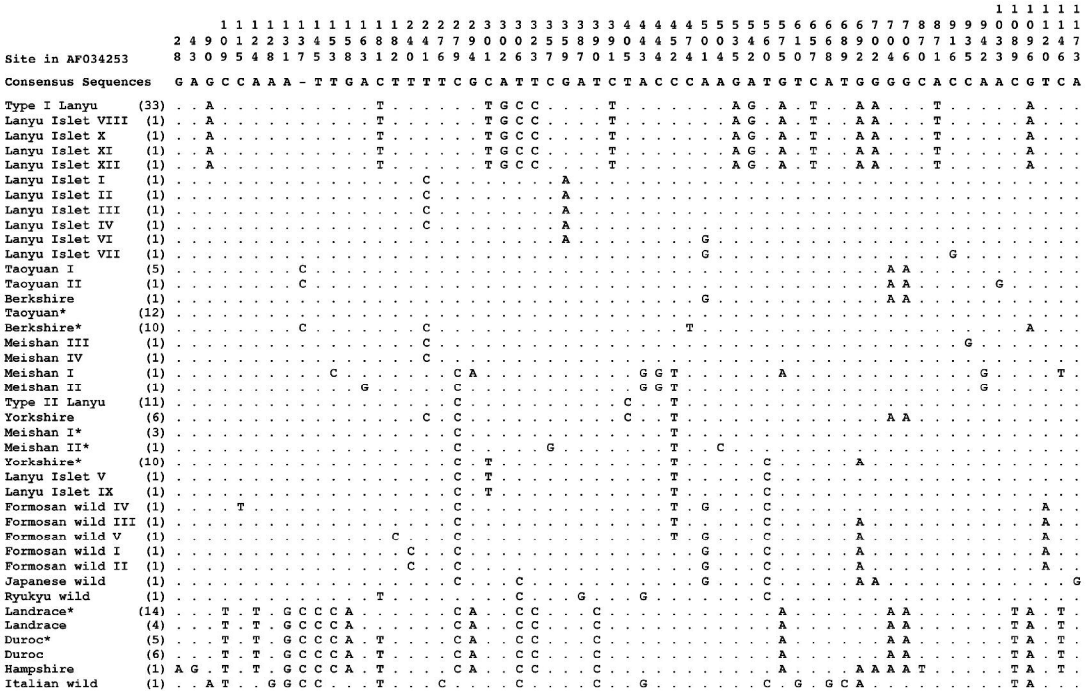


Figure S3. Nucleotide substitutions in mtDNA control regions of conserved Lanyu pigs, Formosan wild pigs, exotic pigs and extant pigs in Lanyu Islet. Sequence breed codes and numbers are given in the first row according to Landrace breed (accession number: AF034253). Only variable sites in mtDNA control region of pigs, with sequence positions given above, are shown. Nucleotides identical to the consensus sequence are denoted by a dot (□), substitution nucleotides different from consensus control region sequence are represented by a different base letter, and deleted nucleotides are denoted by a dash (-). The number inside parentheses in left-hand column indicates the number of pigs used for mtDNA analysis.

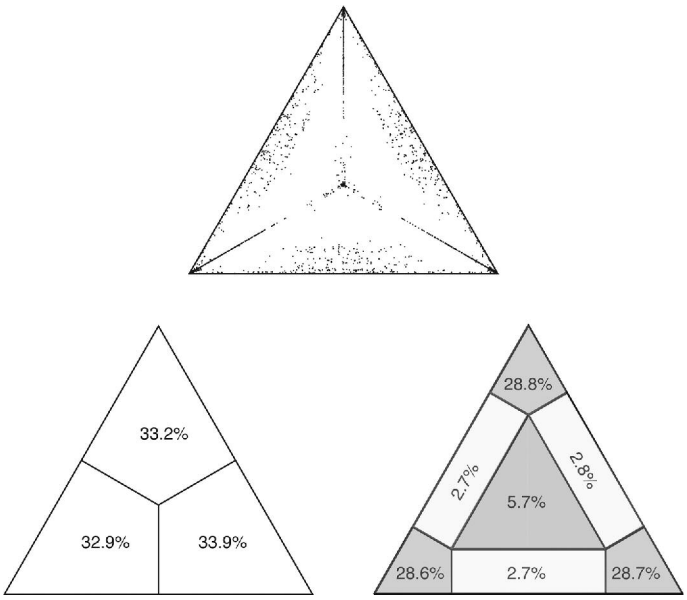


Figure S4. Quartet Puzzling analysis of Maximun Likelihood tree based on the pairwise distance of mtDNA control regions. The treelike topology and phylogenetic signal were obtained by the quartet puzzling method (quartet puzzling support value, 86.1%) supporting the branch assignments in this phylogenetic tree.

Supporting data (Chang et al., the data will submit to Journal of Animal Science on the next month)

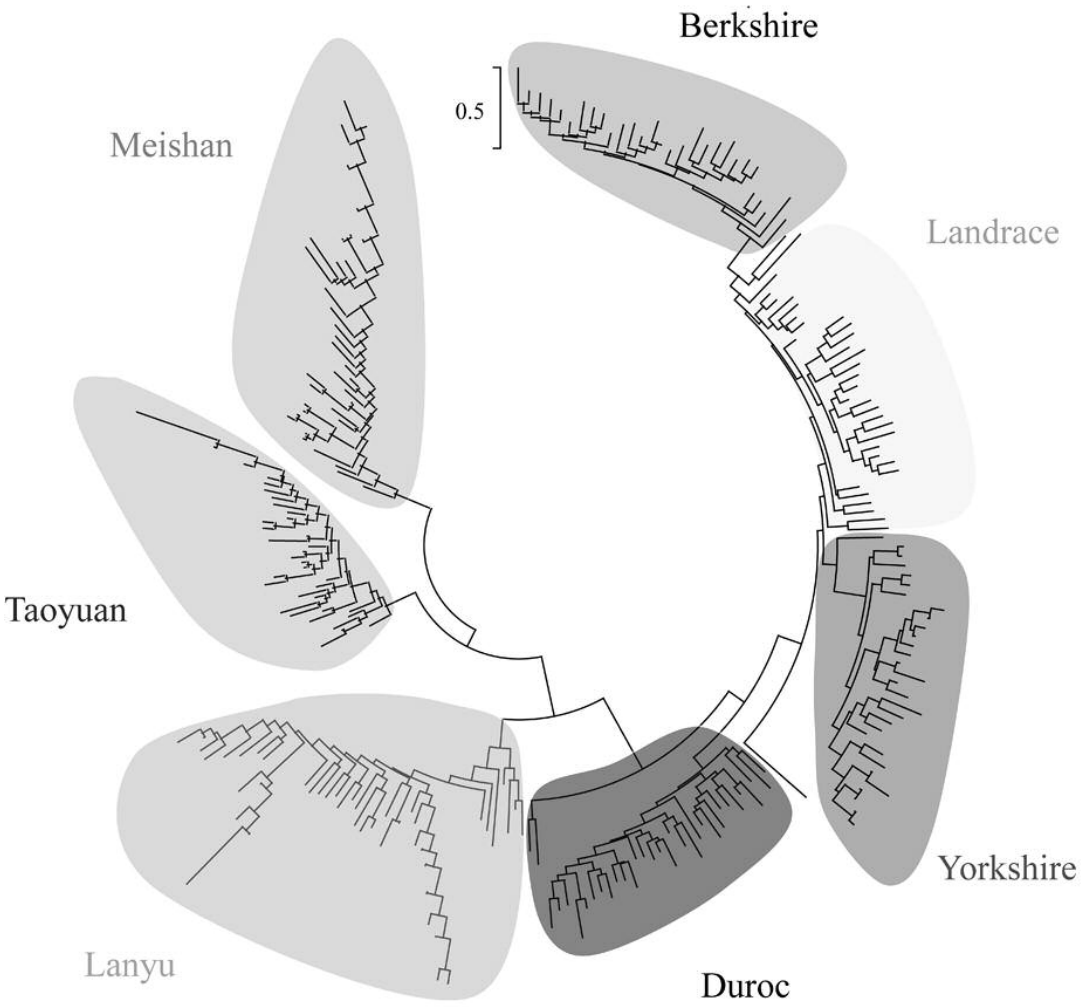


Figure. The Neighbor-Joining tree was constructed based on -ln(proportion of shared allele) distance between 242 individuals of seven pig breeds by 19 microsatellite marker polymorphisms.

	Locations in AF034253	
	Strand	
	H	L
1	432-451	505-484
2	1191-1210	1210-1191
3	1313-1332	1398-1381
4	2199-2216	1806-1787
5	2729-2749	2401-2377
6	3150-3167	2960-2940
7	3693-3713	3278-3259
8	3983-4002	3713-3693
9	4557-4576	4112-4095
10	4754-4771	4846-4829
11	5301-5320	5320-5301
12	5625-5642	5849-5833
13	6201-6222	6222-6201
14	6579-6597	6893-6876
15	7080-7099	7580-7561
16	7561-7580/	7736-7716
17	7629-7646	8231-8212
18	8212-8231	8685-8668
19	8584-8603	9260-9241
20	9241-9260	9568-9587
21	9437-9457	10418-10401
22	10002-10019	10879-10862
23	10515-10535	11584-11571
24	10977-10996	12144-12123
25	11530-11547	12592-12569
26	12123-12144	13381-13361
27	12459-12477	14156-14138
28	13057-13075	114713-14694
29	13605-13625	15057-15038
30	13981-14000	15672-15653
31	14089-14106	15843-15825
32	14694-14713	16581-16564
33	14930-14949	
34	15653-15672	
35	15735-15752	
36	16485-16504	

Table. A list of primers for the full mitochondrial genome sequencing