



**HAL**  
open science

## **Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication**

Hendrik-Jan Megens, Richard P.M.A. Crooijmans, Magali San Cristobal, Xiao Hui, Ning Li, Martien A.M. Groenen

### ► To cite this version:

Hendrik-Jan Megens, Richard P.M.A. Crooijmans, Magali San Cristobal, Xiao Hui, Ning Li, et al.. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. *Genetics Selection Evolution*, 2008, 40 (1), pp.103-128. hal-00894588

**HAL Id: hal-00894588**

**<https://hal.science/hal-00894588>**

Submitted on 11 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication

Hendrik-Jan MEGENS<sup>1\*</sup>, Richard P.M.A. CROOIJMANS<sup>1</sup>,  
Magali SAN CRISTOBAL<sup>2</sup>, Xiao HUI<sup>3</sup>, Ning LI<sup>3</sup>,  
Martien A.M. GROENEN<sup>1</sup>

<sup>1</sup> Wageningen University, Animal Breeding and Genomics Centre, PO Box 338, 6700AH, Wageningen, The Netherlands

<sup>2</sup> INRA, UMR444 Laboratoire de génétique cellulaire, 31326 Castanet Tolosan, France

<sup>3</sup> China Agricultural University, National Laboratories for Agrobiotechnology, Yuanmingyuan West Road 2, Haidian District, 100094 Beijing, P.R. China

(Received 24 July 2006; accepted 18 July 2007)

**Abstract** – Microsatellite diversity in European and Chinese pigs was assessed using a pooled sampling method on 52 European and 46 Chinese pig populations. A Neighbor Joining analysis on genetic distances revealed that European breeds were grouped together and showed little evidence for geographic structure, although a southern European and English group could tentatively be assigned. Populations from international breeds formed breed specific clusters. The Chinese breeds formed a second major group, with the Sino-European synthetic Tia Meslan in-between the two large clusters. Within Chinese breeds, in contrast to the European pigs, a large degree of geographic structure was noted, in line with previous classification schemes for Chinese pigs that were based on morphology and geography. The Northern Chinese breeds were most similar to the European breeds. Although some overlap exists, Chinese breeds showed a higher average degree of heterozygosity and genetic distance compared to European ones. Between breed diversity was even more pronounced and was the highest in the Central Chinese pigs, reflecting the geographically central position in China. Comparing correlations between genetic distance and heterozygosity revealed that China and Europe represent different domestication or breed formation processes. A likely cause is a more diverse wild boar population in Asia, but various other possible contributing factors are discussed.

**pigs / Chinese breeds / European breeds / DNA pools / microsatellite / diversity**

---

\* Corresponding author: hendrik-jan.megens@wur.nl

## 1. INTRODUCTION

Domestication of the pig occurred independently on several occasions [15,25], and some gene flow with the wild boar may have remained until pigs were kept in sties [48]. Historically, there are two major areas of pig breeding, Europe and China, and each seems to represent an independent domestication event. In both areas pig breeding has been applied for at least 7-8000 years [8] resulting in many local breeds but also specialized breeds with a wider distribution, which were selected for the production of particular meat types such as pork or bacon. Europe and China are the origin of about 70% of breed diversity in the world [41] (Europe has 228 listed existing breeds, plus 105 now extinct; China has 118 listed breeds and 10 more extinct). In both China and Europe, the pig has been and remains a major meat producer. However, over the past centuries, pig breeding has shown marked differences between these two areas.

In Europe, pig breeding became decreasingly local, with breeders importing genetic material from elsewhere in Europe [21, 48]. Particularly in the late 18th century many new pig breeds, some the ancestors of currently globally applied commercial breeds, were formed often with the aid of pigs imported from China [8, 48]. In the UK, several popular breeds such as the Berkshire and Large White came into existence and were distributed globally to replace or improve local breeds [21]. Already in 1868, Charles Darwin [10] noted the demise of many indigenous breeds. Nowadays, many of the European local pig breeds have been heavily altered and three-quarters of local or traditional breeds are extinct or marginalized [41].

Conversely, in China traditional local pig breeding has been applied until more recent times [33], although recently these practices have been changing. Although undoubtedly Chinese pig breeders have sought to use animals from other than the local breeds to enhance their breeding practice [12], this has apparently been much less pronounced than in Europe. The result is that China has more pig breeds than any other country in the world by far; the World Watch List for Domestic Animal Diversity [41] lists 118 distinct breeds. Most of these breeds are still local, although unfortunately the advent of modern pig breeding techniques, involving European or new Chinese  $\times$  European synthetic breeds results in rapid marginalization of many of these traditional breeds ([33,41]; 10 listed as endangered, but 30 more show no population data).

There is currently no widely accepted way of grouping European breeds in a classification system. Substantial mixing of types had occurred even before the concept of a 'breed' had been clarified by the end of the 19th century. At the

start of the 20th century, breeds were classified based on production goals (pork or bacon), and in the 1960s a common practice was to divide pigs into “(General) White Meat type”, “Highly Specialised” and “Local Breeds” [33]. Traditional European breeds are sometimes grouped as ‘Celtic’ for Northern Europe vs. Iberian for Southern European breeds. Nowadays, this grouping system is applied mainly for traditional pig breeds from France and northern Iberia vs. the southern “Iberian” breeds [18, 33].

Pig diversity in China has been categorized, based on exterior and production traits and historical data, into six geographic areas [49]. Most local breeds come from the Central Chinese (CC), South Chinese (SC), and South Western Chinese (SWC) areas. The Lower Changjiang River Basin (LCRB) type has perhaps fewer breeds, but is commercially very important as it has some of the most fertile breeds such as Meishan that are currently used to create new synthetic lines. The remaining areas represent the Northern Chinese (NC) pigs, and the Plateau (Plat) type from the Tibetan area, which is the home of a few breeds adapted to marginal feeding and high altitude.

Currently there is a trend to improve and even replace Chinese breeds using European commercial lines, which mirrors the introgression of Chinese breeds into European stock to some extent. A comprehensive, global assessment of pig diversity would give better insight into the genetic relationships and distinctiveness of pigs from different parts of the world, and may aid in preservation of worldwide livestock diversity. Recent papers have reported on relations between pigs using microsatellite markers [12, 22, 26, 28, 40, 47, 50], but mainly between pig breeds from either China or Europe. For instance, Zhang *et al.* [50] concluded that a number of major groups could be discerned, in part congruent with previous classification schemes. By contrast, a similar study [40] showed that for European breeds such subdivisions are very difficult to establish. This discrepancy tantalizingly hints at differences in history and processes of domestication and breed formation and maintenance between these areas. From these studies, it is difficult to gain better insight into differences between Europe and China as there is limited overlap in breeds, and even if overlap existed in markers it would be difficult to reliably merge datasets due to calibration and marker evaluation inconsistencies. The most accurate way to get a better insight into the differences between the Chinese and European breeds would be to include all in a single dedicated study.

A major limitation for global assessment of pig diversity has been the difficulty of obtaining an adequate number of specimens of a reasonable number of breeds from both China and Europe. The current study is part of a large-scale cooperation between many research groups from Europe and China through

the PigBioDiv 1 [32] and PigBioDiv 2 [3] projects, funded by the European Union and the Chinese Academy of Sciences. This cooperation has allowed compilation of a large DNA collection from pig breeds by exchange between laboratories.

A second major limiting factor is cost [38]. A meaningful comparison between pig breeds means genotyping at least dozens of specimens from dozens of breeds from each region to gain adequate insight into diversification patterns, for dozens of markers. For a study such as the present one, this would mean over 150 000 individual genotypings, which would imply an investment in labor and not easily financed consumables. Several meaningful population statistics and comparisons can be deduced from allele frequency data. Although individual genotyping is the most accurate method to estimate allele frequencies in populations, it has been shown that pooled sampling methods [9, 16, 20, 42–44], and references therein can perform very adequately as well, at a fraction of the cost of labor and consumables.

Using DNA pools for 52 populations from Europe and 46 from China, we report on the first comprehensive study on patterns of variation in both of the major areas of pig domestication and breed diversity.

## 2. MATERIAL AND METHODS

### 2.1. Sampling and DNA preparation

A total of 98 lines from Europe and China were sampled; these are listed in Table I including the number of specimens per line used. The material from these lines was collected in the framework of the PigBioDiv 1 [32, 40] and PigBioDiv 2 projects [3] ([www.pigbiodiv2.com](http://www.pigbiodiv2.com)). The European breeds are represented by a number of commercial lines as well as local or regional breeds from essentially the whole of Europe, although pig breeds from the former Soviet Union and Scandinavia were not included. Included are also breeds that originate from the United States, since these breeds such as Hampshire and Duroc are derived in turn (largely) from European pigs. For the current study, 52 European populations out of 58 included in the PigBioDiv 1 project [40] were sampled. The 45 sampled Chinese breeds show a very good representation of all six recognized types [49], including the Tibetan (Plateau) type. Names of European breeds are according to Porter [34], Chinese breeds mainly according to Zhang [49], Zhang *et al.* [50], and Fang *et al.* [12]. DNA was isolated from individual samples using standard phenol-chloroform protocols [32]. Pools of DNA were made by adding equal amounts of DNA for each of the individuals to a single vial for each of the lines [9, 16].

**Table I.** Breeds and lines sampled for this study. PIC = Pig Improvement Company, BHZP = Bundes Hybrid Zucht Program, SCAPAAG = Société Coopérative Agricole pour l'Assainissement et l'Amélioration Génétique du Cheptel Porcin, FH = France Hybrides. For European breeds, it is indicated whether they are considered local (LO), national variety of international breed (I), purebred (C) or synthetic (SY) commercial line. Indicated in the last two columns are Gene Diversity (*H*) and mean number of peaks at > 5% frequency (Np).

<b>Breed</b>	<b>N</b>	<b>Location</b>	<b>Type</b>	<b><i>H</i></b>	<b>Np</b>
Wannan Spotted	39	China, Anhui	CC	0.73	5.62
Xingzi Black	61	China, Anhui	CC	0.70	5.19
Huai	63	China, Fujian	SC	0.70	4.95
Putian	51	China, Fujian	CC	0.69	4.95
Lantang	45	China, Guangdong	SC	0.68	4.95
Bamaxiang	51	China, Guangxi	SC	0.65	4.69
Donglan	52	China, Guangxi	SC	0.75	5.92
Dongshan	53	China, Guangxi	CC	0.67	5.08
Guizhong Small Spotted	51	China, Guangxi	SC	0.62	4.44
Longling	50	China, Guangxi	SC	0.73	5.51
Baixi	35	China, Guizhou	SWC	0.72	5.77
Guanling	47	China, Guizhou	SWC	0.67	4.81
Jiangkou Radish	53	China, Guizhou	SWC	0.73	5.80
Qiandong Spotted	53	China, Guizhou	SWC	0.71	5.54
Xiang	28	China, Guizhou	SC	0.70	5.24
Lingao	52	China, Hainan	SC	0.75	5.92
Wenchang	24	China, Hainan	SC	0.67	4.92
Wuzhishan	48	China, Hainan	SC	0.71	5.03
Min	50	China, Heilongjiang	NC	0.70	4.76
Exi Black	45	China, Hubei	SWC	0.71	5.16
Qingping	47	China, Hubei	CC	0.69	4.86
Tongcheng	50	China, Hubei	CC	0.71	5.26
Yangxing	17	China, Hubei	CC	0.71	5.06
Daweizi	53	China, Hunan	CC	0.74	5.83
Ningxiang	57	China, Hunan	CC	0.70	5.16
Penzhou Mountain	50	China, Hunan	SWC	0.73	5.66
Shaziling	52	China, Hunan	CC	0.70	5.33
Taoyuan	45	China, Hunan	CC	0.75	6.03
Dongchuan	33	China, Jiangsu	LCRB	0.67	4.97
Erhualian	56	China, Jiangsu	LCRB	0.66	4.78
Shawutou	28	China, Jiangsu	LCRB	0.70	4.95
Small Meishan	42	China, Jiangsu	LCRB	0.59	4.00
Dongxiang Spotted	60	China, Jiangxi	CC	0.67	4.68
Leping Spotted	51	China, Jiangxi	CC	0.71	5.38
Nancheng Black	25	China, Jiangxi	CC	0.72	5.35

**Table I.** Continued.

<b>Breed</b>	<b>N</b>	<b>Location</b>	<b>Type</b>	<b>H</b>	<b>Np</b>
Licha Black	54	China, Shandong	NC	0.70	5.32
Hanjiang Black	48	China, Shanxi	NC	0.72	5.64
Mashen	51	China, Shanxi	NC	0.64	4.22
Neijiang	48	China, Sichuan	SWC	0.69	5.03
Yanan	23	China, Sichuan	SWC	0.73	5.36
Tibetan/Zang	39	China, Tibet	PLAT	0.74	5.72
Jiaxing Black	37	China, Zhejiang	LCRB	0.65	4.47
Jinhua	50	China, Zhejiang	CC	0.59	4.03
Shangyu Spotted	50	China, Zhejiang	CC	0.70	5.22
Meishan	14	UK (Roslin)	LCRB	0.64	4.27
Meishan	22	UK (PIC)	LCRB	0.68	4.71
Presticke	50	Czech Republic	LO	0.69	4.87
Créole	50	France	LO	0.65	4.85
Landrace	51	France	I	0.67	4.76
Piétrain	50	France	I	0.63	4.40
Large White	50	France (dam)	I	0.61	4.33
Landrace	51	France (FH)	C	0.63	4.14
Piétrain	48	France (FH)	C	0.62	4.35
Laconie	49	France (Pen Ar Lan)	SY	0.66	4.81
Large White	46	France (Pen Ar Lan)	C	0.67	4.77
Tia Meslan	49	France (Pen Ar Lan)	SY	0.68	4.82
DRB	50	France (SCAPAAG)	SY	0.65	4.53
Large White	34	France (sire)	I	0.59	4.00
Angler Sattelschwein	56	Germany	LO	0.65	4.64
Bunte Bentheimer	47	Germany	LO	0.65	4.57
Large White	52	Germany	I	0.69	5.11
Mangalitsa	29	Germany	LO	0.61	4.00
Piétrain	51	Germany	I	0.64	4.56
Hampshire	60	Germany (BHZP)	C	0.64	4.56
Landrace	55	Germany (BHZP)	C	0.65	4.42
Large White	50	Germany (BHZP)	C	0.63	4.43
Duroc	49	Germany (Schaumann)	C	0.66	4.47
Calabrese	19	Italy	LO	0.60	4.11
Casertana	28	Italy	LO	0.67	4.72
Cinta Senese	30	Italy	LO	0.59	4.06
Duroc	50	Italy	I	0.66	4.70
Landrace	50	Italy	I	0.66	4.85
Large White	50	Italy	I	0.66	4.92
Mora Romagnola	12	Italy	LO	0.51	3.30
Nera Siciliana	50	Italy	LO	0.66	5.00
Pulawska Spot	48	Poland	LO	0.65	4.49

**Table I.** Continued.

<b>Breed</b>	<b>N</b>	<b>Location</b>	<b>Type</b>	<b>H</b>	<b>N<sub>p</sub></b>
Bisaro	60	Portugal	LO	0.69	4.97
Manchado de Jabugo	36	Spain	LO	0.64	4.39
Negro Iberico	48	Spain	LO	0.66	4.62
Negro Canario	18	Spain	LO	0.56	3.70
Retinto	68	Spain	LO	0.67	4.62
Berkshire	50	UK	LO	0.60	4.00
British Lop	35	UK	LO	0.66	4.74
British Saddleback	42	UK	LO	0.68	4.95
Gloucester Old Spot	53	UK	LO	0.59	4.06
Large Black	52	UK	LO	0.62	4.19
Middle White	38	UK	LO	0.64	4.33
Tamworth	42	UK	LO	0.64	4.24
Duroc	50	UK (PIC)	C	0.64	4.41
Hampshire	50	UK (PIC)	C	0.63	4.31
Landrace	50	UK (PIC1)	C	0.65	4.31
Landrace	48	UK (PIC2)	C	0.64	4.44
Landrace	50	UK (PIC3)	C	0.65	4.46
Leicoma	50	UK (PIC)	SY	0.66	4.78
Piétrain	50	UK (PIC)	C	0.66	4.541
Large White	50	UK (PIC1)	C	0.67	4.65
Large White	50	UK (PIC2)	C	0.66	4.53
Large White	50	UK (PIC3)	C	0.64	4.65

## 2.2. Markers and experimental procedures

A total of 39 markers were included in this study that were chosen to maximize the genome coverage. When multiple markers were located in the same chromosome, there was a minimum distance of at least 30 cM to avoid linkage disequilibrium. Markers were further chosen based on (known) absence of null alleles, sharpness of peaks on automated sequencing devices and robustness of PCR [16].

The markers were (in alphanumerical order) the following: CGA, IGF1, S0002, S0025, S0026, S0068, S0070, S0090, S0097, S0101, S0143, S0178, S0215, S0217, S0218, S0225, S0226, S0227, S0301, S0355, SW0024, SW0295, SW0742, SW0787, SW0830, SW0857, SW0902, SW0911, SW0936, SW0951, SW1023, SW1067, SW1111, SW1828, SW2008, SW2406, SW2410, SWR1004, and SWR1941. For further details on markers and PCR conditions for pooled samples see [16].



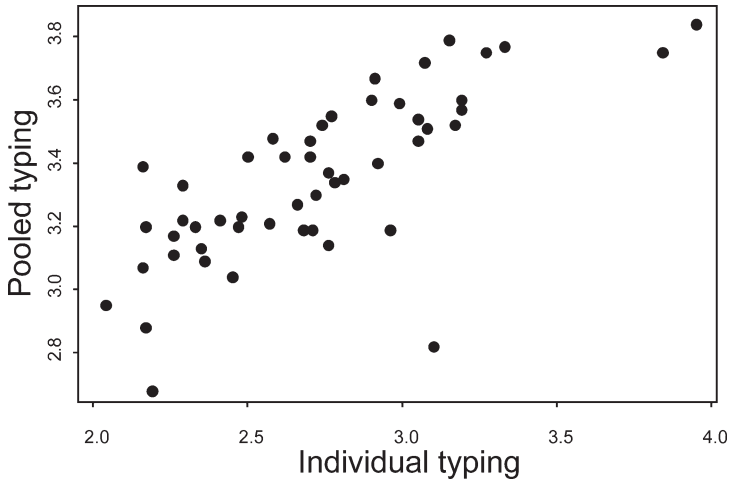
Individual PCR reactions for each of the markers were subsequently pooled into sets of two to five markers in such a way that overlap of alleles was avoided even if markers differed for the fluorescent dye. Sets of markers were analysed on the ABI 3100 capillary automated sequencer, using 36 cm capillaries, and standard microsatellite genotyping settings. Because typing was performed in multiple batches, a calibration panel was used to evaluate inter-batch typing variation.

Instead of allele frequencies, peak frequencies were calculated based on the area under the peaks, using the ABI PRISM<sup>®</sup> GeneMapper<sup>®</sup> Software v3.7. Previous results on pooled chicken DNA [9] revealed that it was more reliable to use peak frequencies rather than to correct for stutter bands. Peak frequencies lower than 5% were discarded and peak frequencies higher than 5% were subsequently re-calculated to add up to 100%.

### 2.3. Data and analysis

In our final data set, we were able to include over 30 markers for all but three lines (Large White sire line from France 22 markers, Guanling 21, Baixi 22), and less than 35 markers were scored in 12 out of 98 breeds. The Phylip program (Phylip 3.64 [14]) does not allow missing data, and therefore custom-written scripting tools (Perl 5.8) were used to perform distance calculations for each pair of populations based on the marker data these populations had in common. All distance calculations were then written to a single distance matrix that was subsequently used for hierarchical clustering using the Neighbor Joining (NJ) procedure [39] as implemented in Phylip. This procedure is conceptually similar to the pair wise gap exclusion options available for DNA sequence distance calculations in programs such as Mega 3.1 [24]. For bootstrapping [13, 14], a similar scripting procedure was implemented.

The distance matrices were used for non-metric Multi Dimensional Scaling (MDS [23]) as implemented in NTSYSpc [37]. MDS was used as an alternative to hierarchical clustering; it is believed to be robust and can be applied in analysis of genetic variation in a geographical context (*e.g.* [2]). Gene diversity ( $H_m$ ) was measured for all loci and all populations by using peak frequencies instead of allele frequencies according to Hillel *et al.* [20]. An average  $H$  was calculated from  $H_m$  for each population across loci, but only for loci that had a higher than 75% success ratio overall.



**Figure 1.** Correlation between individual and pooled typing for effective allele number. The comparison could be made only for European pig populations shared between [40] and the present study.

### 3. RESULTS AND DISCUSSION

#### 3.1. Evaluation of pooled sampling

Pooling DNA samples provides a cost effective method for any study that aims at assessing relative differences of allele frequency among populations [44] and can be applied to microsatellite markers [9, 20, 42] and SNP *e.g.* [46]. Although it is known that pooling DNA samples may lead to a bias in the allele frequency, such estimates can still detect relative differences in allele frequencies among DNA pools [44]. Although its application in pig studies has been reported [16], assessment of its effectiveness is desired. Using peak frequencies rather than correcting for actual alleles results in a systematic overestimation of the actual number of alleles and hence heterozygosity. Nevertheless, as expected [9, 20, 44], the effective number of peaks does correlate strongly with the effective number of alleles as do heterozygosity (Fig. 1) and the inferred genetic distances (see supplementary material online version only: <http://www.gse-journal.org>) when comparing to individually genotyped pigs. Such a comparison was possible only for European breeds, further details can be found in supplementary material. The fact that inferred genetic distances correlate so strongly (supplementary material) ensures that topology of the dendrogram remains the same. Similarly, bootstrap values for pooled samples were congruent albeit somewhat lower than with individual sampling.

Three distance measures available from the Phylip package (Phylip 3.64 [14]) were considered: Nei standard ( $D_s$  [29]), Reynolds ( $D_r$  [36]), and Cavalli-Sforza chord ( $D_c$  [5, 20]). Hillel *et al.* [20] found a significant and large degree of correlation between these three distance measures in chicken lines for pooled samples. Moreover, all distance measures should show a high degree of correlation to the Fst [6]. For most of the analyses in our current paper, we calculated all three to evaluate robustness of the results in lieu of the method applied. Overall, the general conclusions drawn in this paper do not seem considerably biased due to the use of one or the other distance measure, although a few minor discrepancies occur for instance in the topology of the trees.

We chose to report on the results derived from the  $D_c$  measure because first of all it is very similar in nature and performance to the widely used (*e.g.* [2, 12, 50]) Nei  $D_a$  measure [30, 35, 45], and it seems superior in tree building performance particularly for intra-species phylogenies [45].  $D_s$  seems less appropriate for our current study since it cannot cope well with fluctuating populations sizes [14], which is something that is likely to have been common in most breeds. The  $D_c$  and  $D_r$  measures, however, are not designed to allow mutation.  $D_r$  appears to be a good measure for very closely related breeds [27], but may be less appropriate for the current study since the Chinese and European breeds are believed to derive from independent domestication events of wild boar populations [15, 25]. Effectively, breeds from these two regions show a separation in time larger than their domestication histories, as illustrated by the mitochondrial phylogeography [25].

### 3.2. European breeds: clustering

The dendrogram (Fig. 2) is clearly divided in a European and a Chinese group of breeds. In the European part of the dendrogram the different lines of the major commercial breeds tend to group together rather robustly. All Large White lines group together, as do the Landrace, Hampshire and Duroc lines, and in many cases these branches are supported by bootstrap values.

At the base of the European part of the dendrogram is the Duroc. It is known to have originated at least in part from European pigs, but particularly the origin of its' red color has always been questioned. Red European breeds such as the Tamworth or Iberian breeds have been postulated as part of the heritage of the Duroc, although it seems that recorded imports are of too late a date to be able to account for this [33]. Another hypothesis is that the red color is derived from the Red Guinea Hog, which was acquired by slave traders from the coast

of Africa. The fact that the Duroc breed is currently so genetically different from other European breeds adds support for this hypothesis.

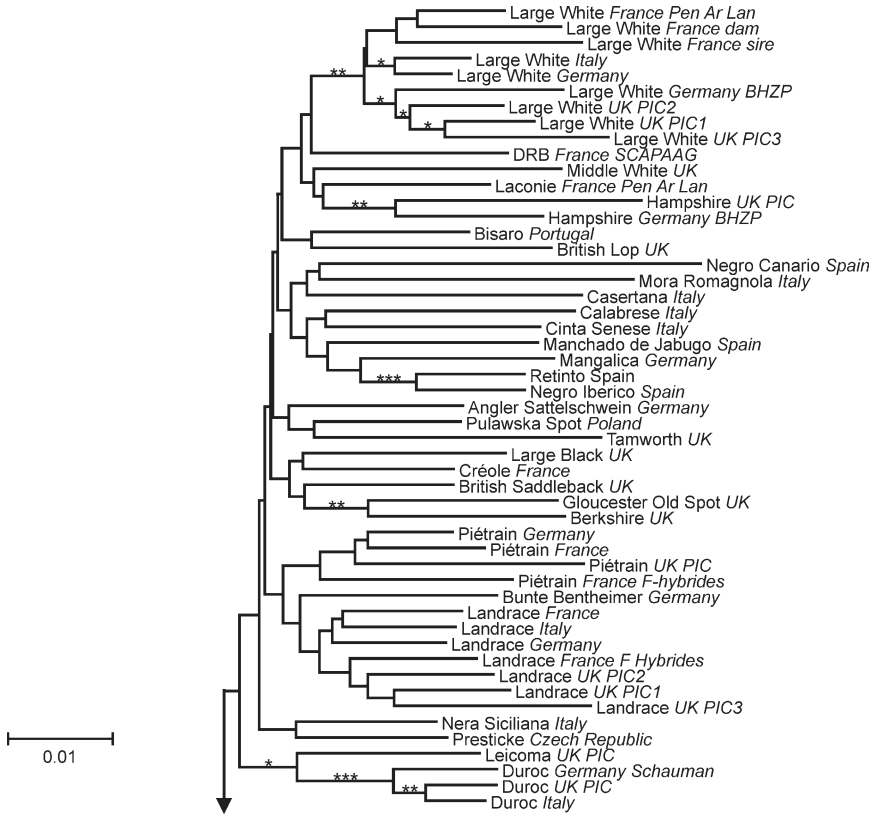
The history of the Pietrain has been largely unknown. Its origin has been thought to lie in French or British prick eared breeds [18], although local landpig has also been postulated [33]. Our results suggest a considerable Landrace heritage for the Pietrain, which is not supported by earlier genetic studies [40]. The Bunte Bentheimer has been crossed with the Pietrain quite extensively [33], and accordingly clusters closely together with this breed.

Historically, it is thought that the Large Black, Berkshire, Gloucester Old Spot and the British Saddleback are part of a south English and Midlands descent, while the British Lop and Large White/Yorkshire are thought to originate from a northern English stock. Furthermore, it is thought that these south/Midlands English groups are all heavily influenced by crossbreeding with Neapolitan and/or Asian pigs in the 18th century [1]. The British Lop is a notable exception of the 'English' group of traditional breeds in Figure 1, which is consistent with the historical records. The Créole on the other hand is part of the 'English' group, and is most closely related to the Large Black. This is in line with the documented history of this uniform black breed, which designates Large Black as the most important founding breed, together with a few others [33].

The Tamworth breed is one of the old English breeds that reportedly has escaped (to some extent) 19th century improvement practices and is believed to have remained more or less in its original state [18, 33]. Fang and Andersson [11] nevertheless found a large proportion of Asian mitochondrial haplotypes in this breed. Although not very dissimilar in overall genetic distance to the Berkshire and Gloucester Old Spot (MDS analysis, results not shown), the Tamworth does also not cluster together with the 'English' group. A 'landrace' origin was postulated by Hammond *et al.* [18], but historically a close relationship with the Berkshire breed is more probable [1, 33]. Interestingly, the Pulawska breed was created in the early 20th century by a cross between Berkshire and local pigs. The fact that it is the Tamworth and not the Berkshire that clusters with this Polish breed may be a result of further crossbreeding and marginalization of the Berkshires [33]. This may have resulted in the Tamworth of today being more like the Berkshire of a century ago than is the present Berkshire breed itself.

The Angler Sattleschwein is also close to the Tamworth and Pulawska. This could be due to the fact that the Pulawska and Sattleschwein share a similar German/Polish landpig heritage, but here more likely to the English connection

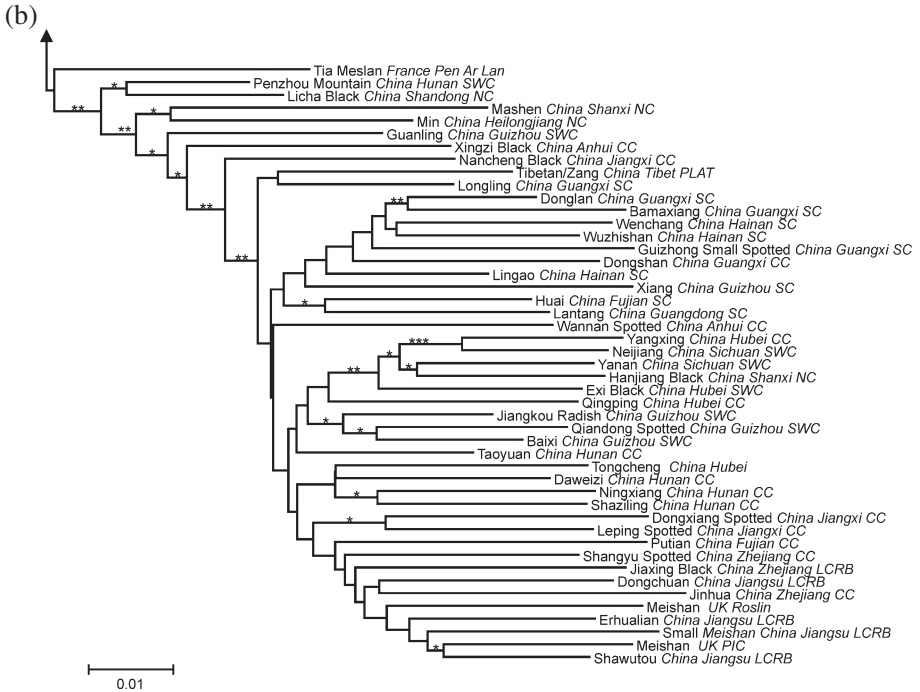
(a)



**Figure 2.** Dendrogram of 98 pig lines from Europe and China, derived from the  $D_c$  using Neighbor Joining clustering. Bootstrap support is indicated at the branches: \* > 50%, \*\* > 70%, \*\*\* > 90%. For Chinese breeds, the types according to Zhang [49] are indicated: NC = North China type, CC = Central China type, SC = South China type, SWC = Southwest China type, Plat = Tibetan/Plateau type, LCRB = Lower Changjiang River Basin type. Figure 2a shows the European part of the dendrogram, Figure 2b the Chinese part. The two parts are connected at the arrow.

that becomes visible; the Angler Sattelschwein has been heavily influenced by the British (Wessex) Saddleback [18].

A southern European cluster (although not supported by bootstrap values) is found in Figure 1, which contains Spanish and Italian breeds, and also the Mangalitsa. Interestingly, furry pigs are thought to have existed for thousands of years [18] and are known from Roman pictures; the Mangalitsa itself is thought to originate from the Balkans [33].



**Figure 2.** Continued.

The grouping of Italian and Iberian breeds actually addresses a long standing question about whether the Italian breeds belong to the same Iberian group as those in the south of Spain, or are intermediates between European and Indochinese pigs. The latter hypothesis has been raised because the black pigs of Italy are believed to be heavily influenced by imported pigs from China – some believe dating back to Roman times although more likely in the 17th century [33]. The black pigs of Italy are of much interest for understanding the history of current commercial breeds because one breed in particular, the Neapolitan, played a pivotal role in the formation of some of the English breeds such as the Berkshire and the Large Black. Although formally extinct, the Neapolitan lineage is thought to live on in similar pigs such as the Casertana and Calabrese. The positioning of the Italian blacks with Iberian breeds that appear to have very little Chinese influences [7, 11], adds support to a shared southern European origin. A recent hypothesis by Larson *et al.* [25] suggests an independent area of domestication on the Italian peninsula opposed to the rest of Europe. Our data does not support this hypothesis.

There are a few exceptions to the ‘southern European’ group, most notably the Nera Siciliana and the Bisaro. These two breeds are now very rare and it is known for both breeds that they have been crossbred heavily in recent times with commercial western European breeds. The Nera Siciliana is listed in the EAAP-Animal Genetic Data Bank ([www.tiho-hannover.de/einricht/zucht/eaap/index.htm](http://www.tiho-hannover.de/einricht/zucht/eaap/index.htm)) as having > 20% introgression of foreign breeds per generation, including Duroc to which it seems rather similar genetically. The Iberian Celtic types to which the Bisaro belongs are thought to be all but extinct and the Bisaro itself has been reported as so heavily crossbred with foreign breeds that it effectively has ceased to exist [33]. The current analysis may be an accurate testimonial to these two breeds’ demise, but may simultaneously also indicate some of the problems associated with evaluating breed history with hierarchical clustering methods; the history of breeds may become so complex that a simple dichotomous scheme may not adequately present the relationships between populations.

Many of the synthetic breeds show interesting genetic relationships that seem perfectly in line with their listed history. The Laconie synthetic, which is reportedly one third each of Hampshire, Piétrain and Large White [33], sits in the dendrogram actually in-between the Hampshires and Large Whites. The ‘ancient’ synthetic breed Middle White clusters are also very close to the Large Whites. The German Leicoma synthetic, based on landraces and local saddlebacks but later heavily crossed with Duroc, shows that the latter part of its heritage is most pronounced by its clustering close to the Duroc lines [33]. The only synthetic line whose history is difficult to reconcile with the position in the tree is the DRB. This line is based on Duroc, although it is noted that there is a lot of introgression from the DRC line, which reportedly is created, next to Duroc, from Large White and Landrace (EAAP-Animal Genetic Data Bank).

Perhaps the most interesting synthetic is the Tia Meslan, which is a cross of Chinese (Meishan × Jiaying) boars and European sows, and has been a closed line since its creation [4]. As a consequence of this split Chinese-European heritage, it is positioned perfectly in-between in the dendrogram.

### **3.3. Chinese breeds: clustering**

Although geographical structure is present in the European breeds to some extent, it is not very pronounced and not well supported in the analysis. This is in contrast with the Chinese breeds. Here we found a very pronounced clustering pattern that largely reflects the types defined by Zhang [49].

The breeds of the North China type are most closely related to the European pigs. This observation was congruent with Fang *et al.* [12] and Kim *et al.* [22] and was in line with the biogeography of the wild boar, that occurs throughout Eurasia but is absent in the desert areas in the Gobi and at high altitudes in the Himalayas [31]. This effectively presented a barrier for dispersion in what is currently the central Asian part of China (*i.e.* Gobi desert and Tibet), and if there was gene flow between Europe and China it was likely through Siberia into Northern China. Phenotypically, the Northern Chinese wild boar is intermediate between European and South Chinese populations [47]. It is thought that throughout history there has been extensive gene flow between domesticated pig and wild boar populations and it is likely that this pattern is still present in geographical diversity in pig populations today.

Perhaps the most commercially interesting group is the extremely prolific Lower Changjiang River Basin (LCRB) type pigs from East Central China. These also include the now 'European' Meishan pigs and together form a rather well defined group in the dendrogram. Included in the LCRB clade is the Jinhua, a breed from the Zhejiang province that is listed as a Central China type (CC) breed. Its grouping with LCRB types reflects the close geographic proximity of this breeds' origin to the LCRB heartland. Within the LCRB cluster, the Taihu pigs, which are sometimes referred to as a single breed [12] (represented here by European Meishan, Small Meishan, Shawutou and Erhualian) form a single group.

A second very distinct group is a South Chinese group (SC), that does not seem to have exceptions apart from the Dongshan, which is listed as a Central Chinese breed. However, Zhang *et al.* [50] also found this breed to be closely related to South Chinese breeds. The Tibetan Zang pig is intermediate between the SC and SWC type pigs. This may reflect its geographic distribution, since the Tibetan actually exists also in Yunnan and Sichuan provinces that are home to SC and SWC type breeds.

The Central Chinese types from Jiangxi and Fujian provinces are positioned exactly in-between LCRB and SC types, whereas the CC types from Hubei and Hunan provinces appear more similar to each other and to the SWC types. This also suggests that genetic relationships are determined in large part by geographic distance [47].

With a few exceptions, the current results for Chinese breeds are congruent with the results obtained by Zhang *et al.* [50]. There are a few breeds that do not group as expected based on either previous studies or based on geography. The Penzhou Mountain appears close to Northern Chinese breeds, which is completely in contrast with previous studies on this breed [12, 50]. The Guanling



breed has previously been reported as clustering firmly together with other SWC type breeds. The Longling pigs do not group as expected with other SC type breeds. The Nancheng Black and Xingzi Black breeds, together in one branch, were expected to cluster with other CC types from Jiangxi province, although Zhang *et al.* [50] also found the Nancheng Black to cluster outside of the other CC types from the same region.

### 3.4. Classification of Chinese pigs

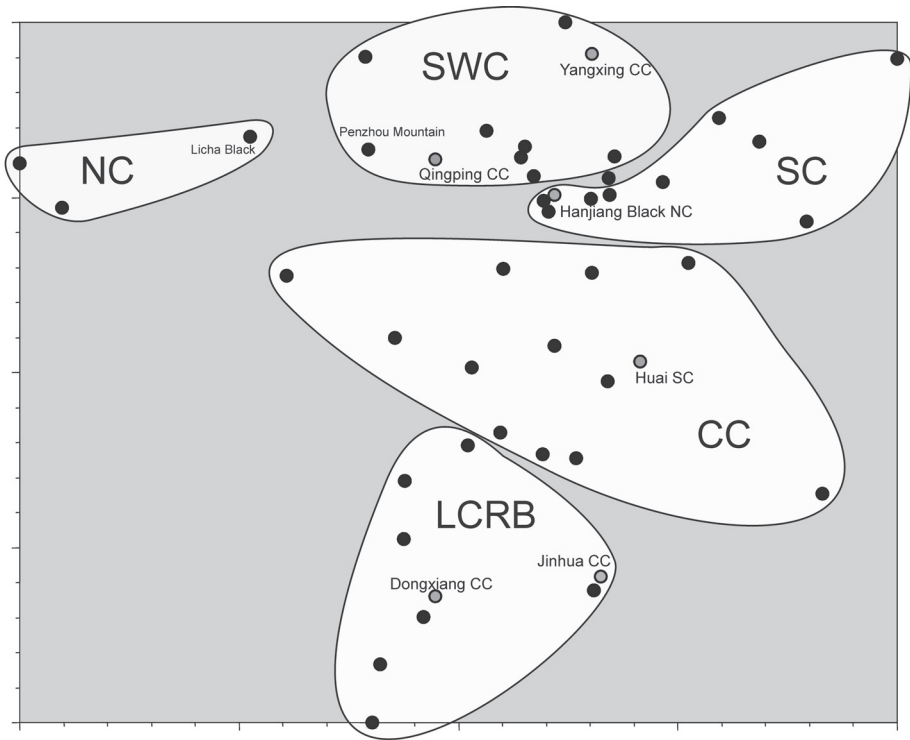
From the study by Zhang *et al.* [50] as well as from our current study there appears to be grounds to revise the classical classification scheme by Zhang [49]. Zhang *et al.* [50] proposed a completely new system based on their phylogenetic analysis, which led to the formation of no less than twelve groups. In this paper, we chose not to adopt this system.

First of all, the classification is based upon a single tree by Zhang *et al.* [50] that is likely to show very poor support at various nodes critical to the classification scheme (but not shown in their study). In our study for instance, the Min and Mashen (and in the MDS analysis also the Licha Black) clearly form a single group, fully congruent to Zhang's [49] northern Chinese type. However, Zhang *et al.* [50] chose to make two groups for Min and Mashen. In our analysis, the Jiaxing Black clusters with the other LCRB types, whereas according to Zhang *et al.* [50] it does not and subsequently is designated as a different group. These examples show that building a classification scheme based on an (partly) unreliable tree will only introduce confusion.

Second, methodologically the groupings, as done by Zhang *et al.* [50], based on 'monophyletic' clusters may seem sound [19], but here we are dealing with populations that show some degree of interbreeding and will, as such, not evolve as discrete entities. For population history studies, especially for those implying a large degree of hybridisation, a strictly dichotomous, hierarchical representation of interrelationships can be a flawed metaphor.

The two-dimensional MDS plot (Fig. 3) shows that some of the inconsistencies with expected clustering in the NJ analysis are in truth artefacts of the clustering algorithm. The Penzhou Mountain is unexpectedly placed in the tree, but in the MDS plot is in fact close to other SWC type breeds. It is also relatively similar to the Licha Black, which in turn is similar to other northern types and as a result of clustering order, the Penzhou Mountain is placed near the base of the Chinese part of the tree. The same principal is true for the unexpected placement of the Guanling and Longling breeds.

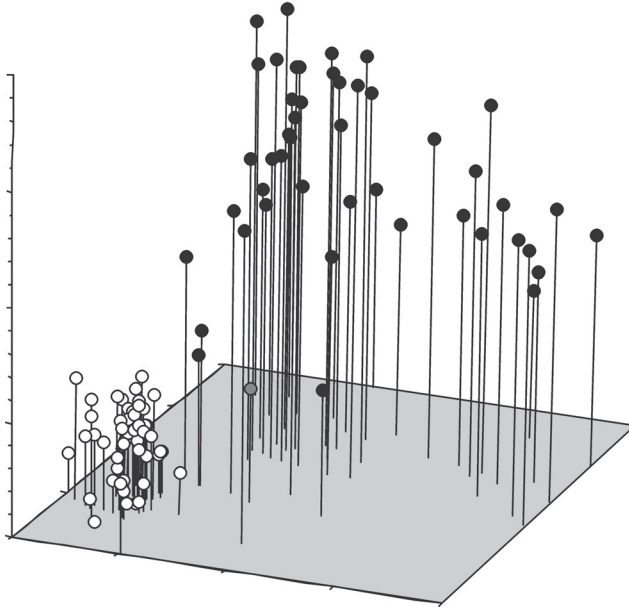
The MDS plot shows a remarkable congruence with the classification scheme by Zhang [49] although a few exceptions exist that for the most part



**Figure 3.** Two-dimensional representation of the Multi Dimensional Scaling (MDS) analysis for Chinese breeds only. Arbitrary lines are drawn around breeds that together form the types recognized by Zhang [49]. Breeds that do not belong to these clusters are indicated with a grey dot and named. See main text for further discussion. NC = North China type, CC = Central China type, SC = South China type, SWC = Southwest China type, LCRB = Lower Changjiang River Basin type. The Tibetan pig is not included in a cluster as it is the only representative of the Plateau type, and is located in this figure between the SWC and SC clusters.

can be attributed to common genetic background due to geographic proximity [12, 47]. The MDS plot can be read virtually as a map of China. The CC types are in the middle of the plot. And then from right to left and from bottom to the top of the graph one goes roughly from north to south geographically.

Accommodating all the exceptions from the tree into small outlier groups creates an unnecessarily complex classification scheme and can only be justified if there is sufficient evidence. Moreover, refining the analysis in a non-hierarchical context actually shows that genetic variation is very much congruent with geography and that there is no need for a new classification scheme.



**Figure 4.** A three-dimensional representation of the Multi Dimensional Scaling (MDS) analysis for all breeds. Chinese breeds have black dots, European breeds have white dots. The synthetic Tia Meslan breed is indicated in grey.

### 3.5. Patterns in variation: Europe vs. China

The dendrogram of Figure 1 shows a somewhat higher degree of bootstrap support for the branches in the Chinese part of the tree compared to the European breeds, and in general there appears to be a far more pronounced geographic structure in genetic variation (Figs. 2 and 3). Both the cluster and MDS analysis for European breeds show almost complete absence of geographic structure (MDS for European breeds not shown).

In addition, the branch lengths of the Chinese part of the tree are longer, which is a result of larger overall inferred genetic distances within this group. A three dimensional representation of the MDS analysis (Fig. 4) shows a marked contrast between Chinese and European breeds. While the European breeds form a single tight cluster, the Chinese breeds are far more spread out. The overall pattern is independent of the distance method used, although it is less pronounced using  $D_r$  and more pronounced with  $D_s$ .

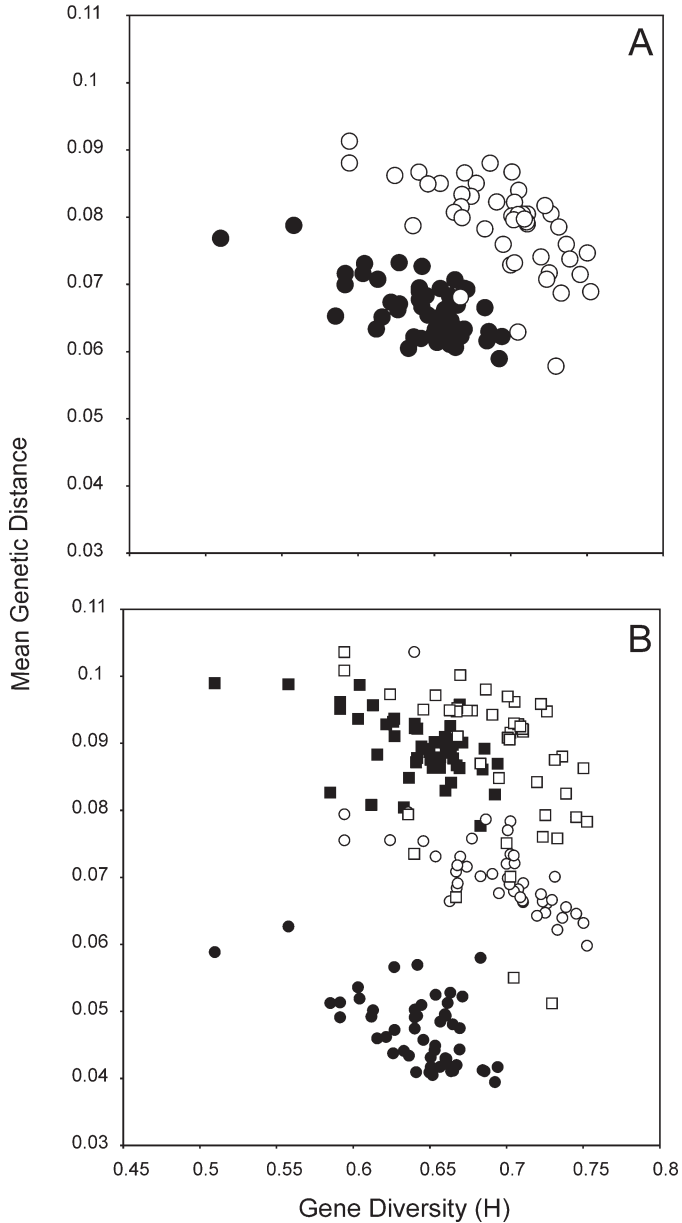
**Table II.** Gene diversity ( $H$ ) and mean number of peaks at the 5% level ( $N_p$ ) for all breeds, European (all) lines, Chinese (all) breeds, European breeds (multiple lines per breed collapsed into a single breed), North Chinese breeds, Central Chinese breeds, Southwest Chinese breeds, South Chinese breeds and Lower Changjiang River Basin breeds.

Region	$H$	$N_p$
All (Europe + China) breeds	0.86	6.05
European lines (all populations)	0.77	4.89
Chinese breeds	0.87	6.24
European breeds	0.77	5.08
South China breeds	0.82	5.76
North China breeds	0.82	6.14
Lower Changjiang Basin breeds	0.81	5.62
Central China breeds	0.85	6.27
Southwest China breeds	0.83	5.92

This contrast in patterns of distances suggests that overall patterns of diversity in pig breeds are different between Europe and China. We further investigated this by plotting gene diversity  $H$  against the Mean Genetic Distance (MGD) according to [20]. MGD is calculated as the mean genetic distance between a given population and all other populations [20], and is similar to  $F_{oi}$  in [40]. A linear relationship between  $H$  and MGD is expected [20, 40]. However, in our analysis we observe not one, but two clear groups that each seems to represent a linear correlation between the two variables (Fig. 5A). What is more important is that these groups largely represent Chinese and European breeds.

This was first observed between Meishan lines and European lines [40]. In their study, it could not be tested whether this was due to the overall larger genetic distance between this particular Chinese breed and all other European lines, or whether this was in fact due to differences between European and Chinese breeds in general. In order to investigate this, we recalculated the MGD first based within regions, meaning European breeds and Chinese breeds, and secondly we recalculated the MGD again, but now based on breeds from the other region. Whichever way we calculate it, the general pattern remains, more pronounced in the former, less pronounced in the latter calculations (Fig. 5B).

Furthermore, gene diversity is higher in China than in Europe (Tab. II), even if we correct for the fact that for Europe we have multiple lines for the important commercial breeds. Conversely, we have also calculated  $H$  of the major regions of pig diversity according to Zhang [49]. Each individual area has



**Figure 5.** A: Mean Genetic Distance (MGD) plotted against Gene Diversity ( $H$ ). Chinese breeds are represented by white dots, European breeds with black dots. B: Modified MGD plotted against Gene Diversity. Chinese breeds are indicated by white symbols, European breeds by black symbols. The circles represent MGD calculated using the same base population, squares represent MGD calculated using the opposite base population (China and Europe).

a larger  $H$  than all European breeds combined. The Central Chinese type has the highest  $H$ , which could be a reflection of its central position.

These differences in patterns of variation could be explained by the relatively recent transformation starting in the late eighteenth century of European pig breeding practices [8, 10, 33, 47]. Such developments included a further shift from free-ranging pigs to sty keeping systems, the introduction of small-boned, prolific and quick maturing pigs from Asia, and an increase in exchange of genetic material throughout Europe to replace or improve local pigs. This transformation was very fast and took less than half a century [8], and as a result many local stocks have gone extinct. However even after this major transformation, local breeds kept on being marginalized or improved to the degree of virtual extinction resulting in a loss of about one-third of the breeds that were recognized in the late 19th century or later [41].

Sty husbandry systems in pig keeping were adopted on a large-scale earlier in China [33], which may have led in already much earlier times to reduction of gene flow with wild populations. Also, China had, until the recent past, a much more restricted degree of movement of pig populations due to legislative and administrative practices, and regional or local breeds appear to have been maintained much more stable. This has probably led to retaining a much more pronounced geographic pattern in genetic variation in China compared to Europe (Figs. 2 and 3)

Individual breeds show a wide range of  $H$  (Tab. I), and substantial overlap exists with breeds from Europe. If Chinese and European breeds were formed out of a single ancestral population and only drift or inbreeding was a major force, we would have expected the lines in Figure 4 to overlap. It is known that pig domestication has been performed independently in China and Europe. Fang and Andersson [11] showed that European mitochondrial haplotypes displayed the signature of a more recent Pleistocene population expansion than the Asian haplotypes. This then would be suggestive of a more diverse wild stock in Asia from which domesticated pigs are derived, which could explain the patterns we observe today.

Both Europe and China have a complex domestication history with multiple local introgressions of wild boar. Furthermore, many cases of introgressions of European with Chinese pigs have been documented. According to some sources [8, 33] these influences have been very large. Asian mitochondrial haplotypes have been found at high proportions in many European pig breeds [7, 11, 15]. One would expect that because of this composite heritage, average  $H$  and allelic diversity could be higher in Europe than in China [17]. Since patterns in microsatellite variation are more congruent with

'ancestral' mitochondrial variation, it seems likely that contributions of sows and boars have not been equal.

Although currently we can only hypothesize on the causes for differences between European and Chinese pig breeds, this study clearly shows that they do represent two different groups not only in general terms of diversity but likely also in domestication and breed formation and maintenance processes. This knowledge should be incorporated in genetic diversity management systems [17, 38]. Crossing pig breeds from the same sub-cluster in the phylogenetic analysis has been proposed as a means of conserving genetic diversity of Chinese pigs [12]. Given the distinctiveness of many Chinese breeds and the larger overall gene diversity compared to European breeds it seems more appropriate to maintain current stocks in sufficient numbers to avoid inbreeding.

### 3.6. Conclusions

The pooled sampling method is a very effective method to discern patterns of relatedness of breeds or populations as shown by our clustering method. Although there are minor differences in details of the trees, in general our results were in line with previous work on Chinese and European breeds [12, 40, 47, 50]. In a few cases the conclusions we draw, are different from earlier published work. This may in part reflect different methods of analysis or differences in interpretation. However, we believe that it also reflects the increased number of markers and breeds included compared to a number of other studies. Most importantly, this study was, for the first time, able to demonstrate patterns of diversity in a large part of the breed diversity of China and Europe simultaneously, demonstrating the differences in genetic variation reflecting differences in domestication and breed formation history between these two major regions of pig domestication.

### ACKNOWLEDGEMENTS

This work was funded by European Union grant QLK5-CT-2002-01059. European DNA samples, other than those provided by the authors, were provided by Agence de la Sélection Porcine, France, Georg-August University Göttingen, Germany, Universidad de Córdoba, Spain, Universidade de Trás-os-Montes e Alto Douro, Portugal, Rare Breeds Survival Trust, UK, Roslin Institute, UK, PIC International Group, UK. We thank Coline van Moorsel for helpful remarks on the manuscript.

**REFERENCES**

- [1] Alderson L., *The chance to survive*, Helm Publishers, Bromley, UK, 1990.
- [2] Beja-Pereira A., Alexandrino P., Bessa I., Carretero Y., Dunner S., Ferrand N., Jordana J., Laloë D., Moazami-Goudarzi K., Sanchez A., Canon J., Genetic characterization of Southwestern European bovine breeds: a historical and biogeographical reassessment with a set of 16 microsatellites, *J. Heredity* 94 (2003) 243–250.
- [3] Blott S., Andersson L., Groenen M., San Cristobal M., Chevalet C., Cardellino R., Li N., Huang L., Li K., Plastow G., Haley C., Characterisation of genetic variation in the pig breeds of China and Europe – the PigBioDiv2 project, *Arch. Zootec.* 52 (2003) 207–217.
- [4] Burlot T., Zhang S., Naveau J., Legault C., Bidanel J.P., Genetic parameters and genetic trends in the Sino-European Tia Meslan composite line, 6th World Congress on Genetics Applied to Livestock Production 23 (1998) 599–602.
- [5] Cavalli-Sforza L.L., Edwards A.W.F., Phylogenetic analysis: models and estimation procedures, *Am. J. Hum. Genet.* 19 (1967) 233–257.
- [6] Cavalli-Sforza L.L., Menozzi P., Piazza A., *The history and geography of human genes*, Princeton University Press, Princeton, NY, USA, 1994.
- [7] Clop A., Amills M., Noguera J.L., Fernandez A., Capote J., Ramon M.M., Kelly L., Kijas J.M.H., Andersson L., Sanchez A., Estimating the frequency of Asian cytochrome B haplotypes in standard European and local Spanish breeds, *Genet. Sel. Evol.* 36 (2004) 97–104.
- [8] Clutton-Brock J., *A natural history of domesticated mammals*, Second Edn. Cambridge University Press, Cambridge UK, 1999.
- [9] Crooijmans R., Groen A.F., van Kampen J.A., van der Beek S., van der Poel J.J., Groenen M.A.M., Microsatellite polymorphism in commercial broiler and layer lines estimated using pooled blood samples, *Poultry Sci.* 75 (1996) 904–909.
- [10] Darwin C., *The variation of animals and plants under domestication*, London, John Murray, 1868.
- [11] Fang M., Andersson L., Mitochondrial diversity in European and Chinese pigs is consistent with population expansions that occurred prior to domestication, *Proc. Biol. Sci.* 273 (2006) 1803–1810.
- [12] Fang M., Hu X., Jiang T., Braunschweig M., Hu L., Du Z., Feng J., Zhang Q., Wu C., Li N., The phylogeny of Chinese indigenous pig breeds inferred from microsatellite markers, *Anim. Genet.* 36 (2005) 7–13.
- [13] Felsenstein J., Confidence limits on phylogenies: an approach using the bootstrap, *Evolution* 39 (1985) 783–791.
- [14] Felsenstein J., PHYLIP (Phylogeny Inference Package) version 3.6, Distributed by the author, Department of Genome Sciences, University of Washington, Seattle, 2005.
- [15] Giuffra E., Kijas J.M.H., Amarger V., Carlborg O., Jeon J.-T., Andersson L., The origin of the domestic pig: independent domestication and subsequent introgression, *Genetics* 154 (2000) 1785–1791.



- [16] Groenen M.A.M., Joosten R., Boscher M.-Y., Amigues Y., Rattink A., Harlizius B., van der Poel J.J., Crooijmans R., The use of microsatellite genotyping for population studies in the pig with individual and pooled samples, *Arch. Zootec.* 52 (2003) 145–155.
- [17] Hall S.J.G., *Livestock Biodiversity. Genetic Resources for the Farming of the Future*, Blackwell Science, Oxford, UK, 2004.
- [18] Hammond J., Johansson I., Haring F., *Handbuch der Tierzucht, Dritter Band Rassenkunde*, Verlag Paul Parey, Hamburg, 1961.
- [19] Hennig W., *Phylogenetic systematics* (English trans. and extensively rev., D. Dwight Davis & Rainer Zangerl), Urbana: U. Ill. Press, 1966.
- [20] Hillel J., Groenen M.A.M., Tixier-Boichard M., Korol A.B., David L., Kirzhner V., Burke T., Barre-Dirie A., Crooijmans R.P.M.A., Elo K., Feldman M.W., Freidlin P.J., Maki-Tanila A., Oortwijn M., Thomson P., Vignal A., Wimmers K., Weigend S., Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools, *Genet. Sel. Evol.* 35 (2003) 533–557.
- [21] Jones G.F., Genetic aspects of domestication, common breeds and their origin, in: Ruvinsky A., Rothschild M.F. (Eds.), *The Genetics of the Pig*, CAB International, Oxon, UK, 1998, pp. 17–50.
- [22] Kim T.H., Kim K.S., Choi B.H., Yoon D.H., Jang G.W., Lee K.T., Chung H.Y., Lee H.Y., Park H.S., Lee J.W., Genetic structure of pig breeds from Korea and China using microsatellite loci analysis, *J. Anim. Sci.* 83 (2005) 2255–2263.
- [23] Kruskal J.B., Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis, *Psychometrika* 29 (1964) 1–27.
- [24] Kumar S., Tamura K., Nei M., MEGA 3: Integrated software for molecular evolutionary genetics analysis and sequence alignment, *Brief. Bioinformatics* 5 (2004) 150–163.
- [25] Larson G., Dobney K., Albarella U., Fang M., Matisoo-Smith E., Robins J., Lowden S., Finlayson H., Brand T., Willerslev E., Rowley-Conwy P., Andersson L., Cooper A., Worldwide phylogeography of wild boar reveals multiple centers of pig domestication, *Science* 307 (2005) 1618–1621.
- [26] Laval G., Iannuccelli N., Legault C., Milan D., Groenen M.A.M., Giuffra E., Andersson L., Nissen P.H., Jorgensen C.B., Beeckman P., Geldermann H., Foulley J.-L., Chevalet C., Ollivier L., Genetic diversity of eleven pig breeds, *Genet. Sel. Evol.* 32 (2000) 187–203.
- [27] Laval G., San Cristobal M., Chevalet C., Measuring genetic distances between breeds: use of some distances in various short term evolution models, *Genet. Sel. Evol.* 34 (2002) 481–507.
- [28] Martinez A.M., Delgado J.V., Rodero A., Vega-Pla J.L., Genetic structure of the Iberian pig breed using microsatellites, *Anim. Genet.* 31 (2000) 295–301.
- [29] Nei M., Genetic distance between populations, *Amer. Nat.* 106 (1972) 283–292.
- [30] Nei M., Tajima F., Tateno T., Accuracy of estimated phylogenetic trees from molecular data, *J. Mol. Evol.* 19 (1983) 153–170.
- [31] Oliver W.L.R., Brisbin I.L., Takahashi S., The Eurasian wild pig (*Sus scrofa*), in: *Pigs, Peccaries and Hippos: Status Survey and Action Plan*, W.L.R. Oliver (Ed.), Gland, Switzerland: IUCN, 1993, pp. 112–121.

- [32] Ollivier L., Amigues Y., Boscher M.-Y., An EC-funded project on characterization of genetic variation in the European pig. Objectives, organization, breed sampling, DNA preparation and circulation, *Arch. Zootec.* 52 (2003) 137–144.
- [33] Porter V., *Pigs. A Handbook to the Breeds of the World*, Helm Information Ltd., Near Robertsbridge, UK, 1993.
- [34] Porter V., *Mason's World Dictionary of Livestock Breeds, Types and Varieties*, 5th edn., CABI Publishing, Wallingford, UK, 2002.
- [35] Reif J.C., Melchinger A.E., Frisch M., Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management, *Crop Sci.* 45 (2005) 1–7.
- [36] Reynolds J.B., Weir B.S., Cockerham C.C., Estimation of the coancestry coefficient: basis for a short-term genetic distance, *Genetics* 105 (1983) 767–779.
- [37] Rohlf F.J., *NTSYSpc 2.1, Numerical Taxonomy and Multivariate Analysis System*, Exeter Software, Setauket, NY, USA, 2000.
- [38] Ruane J., A critical review of the value of genetic distance studies in conservation of animal genetic resources, *J. Anim. Breed. Genet.* 116 (1999) 317–323.
- [39] Saitou N., Nei M., The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol.* 4 (1987) 406–425.
- [40] San Cristobal M., Chevalet C., Haley C.S., Joosten R., Rattink A.P., Harlizius B., Groenen M.A.M., Amigues Y., Boscher M.-Y., Russell G., Law A., Davoli R., Russo V., Désautés C., Alderson L., Fimland E., Bagga M., Delgado J.V., Vega-Pla J.L., Martinez A.M., Ramos M., Glodek P., Meyer J.N., Gandini G.C., Matassino D., Plastow G.S., Siggens K.W., Laval G., Archibald A.L., Milan D., Hammond K., Cardellino R., Genetic diversity within and between European pig breeds using microsatellite markers, *Anim. Genet.* 37 (2006) 189–198.
- [41] Scherf B.D., *World Watch List for Animal Diversity*, 3rd edn., FAO, Rome, Italy, 2000.
- [42] Schnack H.G., Bakker S.C., Van't Slot R., Groot B.M., Sinke R.J., Kahn R.S., Pearson P.L., Accurate determination of microsatellite allele frequencies in pooled DNA samples, *Eur. J. Hum. Genet.* 2 (2004) 925–934.
- [43] Silva L.K., Liu S., Blanton R.E., Microsatellite analysis of pooled *Schistosoma mansoni* DNA: an approach for studies of parasite populations, *Parasitology* 132 (2006) 1–8.
- [44] Skalski G.T., Couch C.R., Garber A.F., Weir B.S., Sullivan C.V., Evaluation of DNA Pooling for the estimation of microsatellite allele frequencies: a case study using Striped Bass (*Morone saxatilis*), *Genetics* 178 (2006) 863–875.
- [45] Takezaki N., Nei M., Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA, *Genetics* 144 (1996) 389–399.
- [46] Yang H.-C., Liang Y.-J., Huang M.-C., Li L.-H., Lin C.-H., Wu J.-Y., Chen Y.-T., Fann C.S.J., A genome-wide study of preferential amplification/hybridization in micro-array-based pooled DNA experiments, *Nucleic Acids Res.* 34 (2006) 1–11.
- [47] Yang S.-L., Wang Z.-G., Liu B., Zhang G.-X., Zhao S.-H., Yu M., Fan B., Li M.-H., Xiong T.A., Li K., Genetic variation and relationships of eighteen Chinese indigenous pig breeds, *Genet. Sel. Evol.* 35 (2003) 657–671.

- [48] Zeuner F.E., *A History of Domesticated Animals*, Harper and Row Publishers, New York, NY, USA, 1963.
- [49] Zhang Z., *Pig Breeds in China*, Shanghai Scientific and Technical Publishers, Shanghai, China, 1986.
- [50] Zhang G.-X., Wang Z.-G., Sun F.-Z., Chen W.-S., Yang G.-Y., Guo S.-J., Li Y.-J., Zhao X.-L., Zhang Y., Sun J., Fan B., Yang S.-L., Li K., Genetic diversity of microsatellite loci in fifty-six Chinese native pig breeds, *Acta Genet. Sinica* 30 (2003) 225–233.