

Microsatellite-based characterization of Southern African domestic pigs (*Sus scrofa domestica*)

H. Swart^{1,2}, A. Kotze^{3,4}, P.A.S. Olivier¹ and J.P. Grobler^{3#}

¹Department of Biodiversity, University of Limpopo, P/Bag X1106, Sovenga 0272, South Africa

²Animal Improvement Institute, Agricultural Research Council, P/Bag X2, Irene 0062, South Africa

³Department of Genetics, University of the Free State, PO Box 339, Bloemfontein 9300, South Africa

⁴National Zoological Gardens of South Africa, PO Box 754, Tshwane 0001, South Africa

Abstract

This paper details genetic characterization and trends from a microsatellite-based study of genetic diversity on southern African pig populations. A total of 351 pigs from three commercial breeds and three indigenous populations were genotyped at 39 loci. Differences among the levels of genetic diversity in populations correlated well with known population histories. In commercial breeds, heterozygosity was higher in the well established SA Landrace and Large White breeds (0.580 and 0.636) compared to the Duroc breed, established more recently (0.531). In indigenous populations, the highest heterozygosity levels were found in the Mozambican and South African populations (0.692 and 0.634) with a lower value of 0.531 in a smaller Namibian population. A hierarchical division of total genetic diversity revealed a high between-population component of 17.9%. F_{ST} - and R_{ST} -based analysis confirmed high levels of differentiation, with pair-wise comparisons between breeds indicating significant differentiation in 20 out of 21 comparisons. Results from an assignment test confirmed results from F_{ST} and R_{ST} and suggested a true genetic structure with significant differentiation between most populations sampled, but with little differentiation among the commercial SA Landrace and Large White breeds. The results are discussed with reference to known historical information on commercial and indigenous pig populations. This paper also presents new data on the optimization of microsatellite markers for application in *Sus scrofa domestica*.

Keywords: Genetic diversity, domestic pig, microsatellites, indigenous breeds

Corresponding author. E-mail: groblerjp@ufs.ac.za

Introduction

The overall southern African pig population consists of a mixture of pigs introduced through diverse pathways. Migration between the 3rd and 7th centuries brought Iron Age communities with domesticated animals into the eastern parts of South Africa, with one archaeological record of a pig introduction at this time (Plug & Badenhorst, 2001). However, pigs most likely played an insignificant role as livestock. Early pastoralists in southern Africa found them unsuitable for a nomadic lifestyle while religious taboos, diseases and habitat favoured alternative types of livestock (Briggs, 1983; Plug, 1993; Bester & Küsel, 1998). Consequently pig remains are not common in southern African excavation sites (Epstein & Mason, 1971). In more recent centuries, pigs were bartered from Chinese trading ships passing South African shores (Ramsay *et al.*, 1994). All other records from the sub-region date to post-European contact (Plug & Badenhorst, 2001). Passing Portuguese travellers imported pigs as a medium of exchange with local populations (Quin, 1959), and a last phase of introduction began in the 16th to 17th century when European breed types were introduced (Bester & Küsel, 1998).

According to figures from the South African Department of Agriculture, Forestry and Fisheries, pig farming in South Africa had developed into a national industry with an estimated 1,613 million pigs by 2009 (DAS, 2010). Modern methods are used for performing recording, including the PIG BLUP system applied by the ARC Animal Improvement Institute from 1993 (Visser, 2004). Established commercial breeds farmed in South Africa include the South African Landrace, the Large White and the Duroc. Ancestors of the South African Landrace were originally developed in Denmark and from 1949 exported to various countries from 1949 (Treacy, 1976). It is now the second best represented commercial breed in the country (Visser *et al.*, 1993). The Large White was first recognized as a distinct breed in England in 1884 (Briggs, 1983). This breed is currently one of the two major maternal populations in the world (Ruvinsky & Rothschild,

1998). The Duroc was developed in the USA and has become important as terminal sires in many countries (Ruvinsky & Rothschild, 1998). It is the fastest growing breed in the world during the last quarter of a century. The first Duroc pigs have been imported into South Africa from Canada in 1980 primarily for cross-breeding purposes (Visser *et al.*, 1993). Molecular data describing levels of diversity in these breeds following various periods of utilization in South Africa are not available.

A number of distinct indigenous pig populations have been identified in southern Africa, but historical information on the origins of these populations is very limited. The Kolbroek is an indigenous South African phenotype of unknown origin. Kolbroek pigs are hardy, survive by scavenging outside homesteads, do well on a high fibre diet and have high disease tolerance and a docile nature (Visser, 2004). Namibia also hosts a population of indigenous pigs of uncertain origin. In Mozambique, pigs of an unimproved type are found around villages, for example in the Tete Province.

Indigenous pigs were long regarded unsuitable for intensive commercial breeding because of their slow growth and inadequate meat production (Prolit, 2004). Nevertheless, many local African domestic animal populations exhibit well-established adaptations to prevailing environmental and management conditions. This represents a valuable genetic resource for improving breeds and the efficiency of animal production. Indigenous populations may add potential characteristic genetic variation which could be valuable to producers in supplying new diversity for the improvement of commercial lines (Blott *et al.*, 2003). Furthermore, a progressive trend towards more environmentally effective pig production and small-farm systems, together with data on the potential of indigenous pigs as converters of fibre, root and leaf crops, have shown that these animals are invaluable to the small-scale rural farmer (Ramsay *et al.*, 1994).

The status of genetic diversity in southern African indigenous and commercial pig populations is currently unknown. This study aimed to genetically characterize the three principal commercial pig populations of South Africa and the indigenous pig populations of southern Africa, and to determine levels of diversity within specific breeds and populations, using molecular markers.

Materials and Methods

Commercial pig breeds sampled, were the SA Landrace (SAL), Large White (LAW) and Duroc (DUR), with indigenous populations from specific regions represented by Namibia (NAM), Mozambique (MOZ) and South Africa (Kolbroek - KOL). The term population was used in preference to breed for the indigenous pigs. Sample sizes and specific origins of pigs are presented in Tables 1 and 3. In order to obtain the most unbiased sample for each nominal breed or national population, an approach was followed in which samples were received from a variety of sources and from different geographical regions. The Kune-Kune (KK) breed, an indigenous breed from New Zealand, was selected as the outgroup.

Hair samples with visible roots were collected from identified populations. The extraction of genomic DNA from hair roots was based on the method of Higuchi *et al.* (1988). A panel of 39 microsatellite loci distributed throughout the genome and identified as part of the Pig Genome Project was used in this study. The markers used (with chromosome position) were: S0073 (4), SW35 (4), S0298 (16), SW1134 (5), SW1851 (1), SW2456 (X/Y), SW2514 (2), SW983 (9), S0120 (18), SW2 (5), SW1557 (14), SW378 (5), SW761 (14), SW2008 (11), SW995 (5), SW352 (7), SW472 (7), SW949 (X/Y), S0004 (15), S0165 (3), S0217 (4), SW225 (13), SW1041 (10), SW21 (9), SW2404 (4), S0035 (6), S0006 (16), SW749 (9), SW2410 (8), SW940 (9), S0295 (9), SW839 (4), SW2406 (6), SW2419 (6), SW316 (6), S0121 (7), SW322 (6), S0385 (11) and SW2443 (2) (Fredholm *et al.*, 1993; Ellegren *et al.*, 1994; Hoyheim *et al.*, 1994; McQueen *et al.*, 1994; Robic *et al.*, 1994; Rohrer *et al.*, 1994; Groenen *et al.*, 1995; Alexander *et al.*, 1996; Rohrer *et al.*, 1996; Lopez-Corrales *et al.*, 1999).

Each marker was first amplified separately, and the results analyzed to determine fragment size ranges and the correct primer concentration, annealing temperature and volume of DNA for each primer set. Similarly labelled primers with overlapping fragment sizes were then assigned to separate PCR multiplexes. Initially, PCR mixtures contained deionized water, 25 μ M dNTP's, 0.3 mM 10 x Supertherm Gold™ reaction buffer (20 mM Tris-HCl pH 8.3, 15 mM MgCl₂, 50 mM KCl), Supertherm Gold™ DNA polymerase (2.5 U), different primer concentrations and 20 ng extracted genomic DNA, in a total volume of 8.2 μ L. The final PCR protocol was optimized with different primer concentrations for each multiplex. Reaction conditions consisted of a 10 min Hot Start® polymerase activation step at 95 °C, 35 cycles of 45 sec

denaturation at 94 °C, 60 sec annealing at 60 °C, 60 sec extension at 72 °C, and a final extension step at 72 °C for 60 min.

Fluorescently labelled DNA fragments were separated using an ABI377 automated sequencer. Polyacrylamide gels were prepared and samples loaded in accordance to the manufacturer's instructions. Fragment analysis was carried out using GeneScan™ and GenotyperII™ software. The Hex® 350 size standard was used as reference to ensure the correct assignment of each sample.

Table 1 The number of pigs sampled and specific origin for commercial breeds and indigenous populations

Breed	Sampling sites	Number of Samples
Indigenous Namibia	Rundu, Kavango region	24
Indigenous Mozambique	Angonia, Tete Province	166
Indigenous South Africa (Kolbroek)	Irene, Gauteng	33
	Eastern Cape	9
	Western Cape	4
	Sasolburg, Gauteng	1
	Chrissiesmeer, Mpumalanga	7
	Rouxville, Free State	6
	Hoekville, Northern Cape	7
	Mosselbay, Western Cape	8
Large White	Stellenbosch, Western Cape	5
	Baynesfield, KwaZulu-Natal	8
	Pretoria, Gauteng	3
	Howick, KwaZulu-Natal	8
	Magaliesburg, North West	4
	Somerset-West, Western Cape	3
SA Landrace	Stellenbosch, Western Cape	5
	Howick, KwaZulu-Natal	8
	Mooi River, KwaZulu-Natal	3
	Polokwane, Limpopo	3
	Magaliesburg, North West	3
	Somerset-West, Western Cape	3
Duroc	Stellenbosch, Western Cape	4
	Howick, KwaZulu-Natal	8
	Mooi River, KwaZulu-Natal	3
	Polokwane, Limpopo	3
	Magaliesburg, North West	1
	Somerset-West, Western Cape	3
Kune-Kune	New Zealand	21
Total		351

A MS Toolkit (Park, 2001) was used to calculate allele frequencies and genetic diversities within breeds and to prepare input files for the other software used. Alleles were classified as private alleles based on a threshold frequency of 0.05; following Budowle *et al.* (1996). Genetic diversity within populations was quantified using measures that reflected overall levels of genetic diversity within nominal populations of each breed and indigenous population rather than all individual local populations. It was thus not considered valid to test for deviations from HWE or for the rate of inbreeding (F_{IS}), because the overall samples per

Table 2 Primers per multiplex, fluorescent labels used, allele size ranges and primer concentrations for seven multiplex reaction sets

	Label	RANGE	pMol Forward	pMol Reverse
PLEX 1				
S0073	Fam	90 - 125	5	5
SW35	Fam	130 - 170	7	7
S0298	Tet	170 - 180	5	5
SW1134	Tet	130 - 145	8	8
SW1851	Tet	85 - 105	12	12
SW2456	Hex	195 - 210	7.5	7.5
SW2514	Hex	105 - 135	12	12
SW983	Hex	75 - 100	5	5
PLEX 2				
S0120	Fam	145 - 180	6	6
SW2	Fam	90 - 130	5	5
SW1557	Tet	85 - 110	10	10
SW378	Tet	120 - 130	9	9
SW761	Tet	150 - 170	5	5
SW2008	Hex	90 - 110	8	8
PLEX 3				
SW995	Fam	150 - 170	6	9
SW352	Fam	110 - 115	10	10
SW472	Tet	88 - 102	6	6
SW949	Tet	180 - 215	6	6
S0004	Hex	160 - 175	8	8
PLEX 4				
SO165	Fam	135 - 170	5	5
SO217	Tet	140 - 160	4	4
SW225	Tet	90 - 115	7	7
SW1041	Hex	95 - 106	4	4
SW21	Hex	120 - 145	5	5
SW2404	Hex	160 - 190	7	7
PLEX 5				
S0035	Fam	160 - 200	6	6
S0006	Tet	235 - 250	5	5
SW749	Tet	102 - 116	6	6
SW2410	Hex	100 - 125	4	4
SW940	Hex	145 - 160	10	10
PLEX 6				
S0295	Fam	225 - 270	6	6
SW839	Fam	145 - 180	5	5
SW2406	Tet	220 - 230	7.5	7.5
SW2419	Tet	120 - 140	10	10
SW316	Hex	150 - 170	10	10
PLEX 7				
S0121	Fam	215 - 270	12	12
SW322	Tet	110 - 122	7	7
SW0385	Hex	140 - 155	8	8
SW2443	Hex	200 - 215	12	7

breed would show evidence of the Wahlund effect due to the number of sub-populations sampled, whereas sample sizes for individual local populations were too small for statistical significance. Measures used were Nei's unbiased heterozygosity (H_z , Nei, 1987), observed heterozygosity (H_o) and the average number of alleles per locus, all from the MS Toolkit. Since the observed number of alleles in a sample is highly dependent on sample size (Rodrig  n  z *et al.*, 2008), we also calculated allelic richness (R_s), which takes into account the variation in sample size (El Mousadik & Petit, 1996; Petit *et al.*, 1998; Toro *et al.*, 2008). The expected number of alleles in a sub-sample was thus calculated using FSTAT software (Goudet, 2001).

Analysis of Molecular Variance (AMOVA) as implemented in ARLEQUIN software (Excoffier *et al.*, 2005) was used to yield estimations of population structure at different levels of a specified hierarchy. As an additional measure of genetic differentiation between populations, pair-wise F_{ST} and associated p values were calculated among populations, using the ARLEQUIN programme. We also calculated among population R_{ST} values as a supplement to F_{ST} to specifically cater for the stepwise mutation model through which new microsatellite alleles are created, using RSTCALC software (Goodman, 1997). The Bonferroni correction for multiple comparisons (Rice, 1989) was applied to all p values from F_{ST} and R_{ST} estimates to compensate for possible type I errors resulting from multiple pair-wise comparisons.

A Bayesian-based method (Pritchard *et al.*, 2000) was used to infer the true number of genetic populations (clusters or K) in the dataset. Prior data on population boundaries were thus ignored before testing for and identifying distinct genetic populations, and assigning individuals to populations. The STRUCTURE programme (Pritchard *et al.*, 2000; Falush *et al.*, 2003) was used to implement a Bayesian-based assignment approach for this analysis. The model used for simulation was based on an assumption of admixed ancestry and correlated allele frequencies. To estimate the true number of populations the parameter $\text{Pr}(X | K)$ was applied, where K is a value of 1 - 7. Thirteen independent runs for each K were used. All runs were carried out with a burn-in period of 20,000 steps followed by 100,000 MCMC (Markov Chain Monte Carlo) iterations (Pritchard *et al.*, 2000). The probability value for each K, averaged over 13 runs, and the standard deviation (SD) for 13 runs of each K, was then determined. The simulation was then run again for the most likely K, with a burn-in of 100,000 steps followed by 1,000,000 MCMC repeats, to elucidate the real distribution of pigs from each population to the identified clusters.

Results

Results from optimization suggested that 5 - 9 hairs (depending on coarseness) were needed to provide sufficient template DNA for a PCR reaction. The 39 primers used were assigned to seven PCR multiplexes based on dye label, fragment size ranges, primer concentration required and annealing temperature (Table 2). Artefacts of multiplexing such as allelic dropout and the amplification of multi-allelic peaks were minimal.

A total of 445 alleles were detected in 39 microsatellite loci screened, including 122 private alleles (Table 3). The number of private alleles in individual groups ranged from 5 - 45, and these were observed in all populations except pigs from Namibia. The unbiased heterozygosity (H_z) estimates are presented in Table 3. Values of H_z ranged from 0.531 - 0.636 in commercial populations, and 0.531 - 0.692 in indigenous populations, with an average of 0.611 across all populations. Observed heterozygosity (H_o) estimates are also presented in Table 3. The average allele number (across populations) was 11.13. Within individual commercial populations, values ranged from 3.98 - 5.50, with a broader range of 3.93 - 8.45 in indigenous populations (Table 3). Following correction for sample size, allelic richness was calculated based on a minimum sample size of 13 individuals per population (Table 3). Values of R_s ranged from 3.65 - 4.64 in commercial populations and 3.58 - 5.46 in indigenous populations.

Results from AMOVA indicated that 82.1% of the total genetic variation was caused by differences within populations, 13.8% by differences among populations within designated groups (either commercial or indigenous) and 4.1% between commercial and indigenous groups. Values indicating the magnitude of divergence between all population pairs are presented in Table 4. All F_{ST} and R_{ST} P-values support the hypothesis of significant ($P < 0.05$) differentiation between population pairs after Bonferroni correction, except for the combination of Mozambique and Namibia. F_{ST} values among commercial populations and among indigenous populations were 0.112 - 0.189 and 0.123 - 0.204, respectively, and generally higher at 0.128 - 0.270 for pair-wise combinations between commercial and indigenous populations. Divergence between the Mozambican and Namibian pig populations was, however, low. The trends observed for F_{ST} also apply to R_{ST} .

Table 3 Sample size (n), number of private alleles, unbiased heterozygosity (Hz), observed heterozygosity (Ho), average number of alleles per locus (A) and allelic richness (Rs) in seven pig populations

	SAL	LAW	DUR	NAM	MOZ	KOL	KK
n:	26	31	22	24	166	61	21
Private alleles:	6	5	8	0	45	16	42
Hz:	0.580 (± 0.032)	0.636 (± 0.021)	0.531 (± 0.035)	0.531 (± 0.035)	0.692 (± 0.023)	0.634 (± 0.024)	0.675 (± 0.036)
Ho:	0.522 (± 0.016)	0.584 (± 0.014)	0.504 (± 0.017)	0.518 (± 0.016)	0.609 (± 0.006)	0.537 (± 0.010)	0.508 (± 0.020)
A:	5.48	5.50	3.98	3.93	8.45	6.18	5.97
Rs:	4.64	4.59	3.65	3.58	5.46	4.80	4.80

SAL - SA Landrace; LAW - Large White; DUR - Duroc; NAM – Namibia; MOZ – Mozambique; KOL – Kolbroek; KK - Kune-Kune.

Table 4 F_{ST} values (below the diagonal) and R_{ST} values (above the diagonal) among seven pig populations. All pairwise comparisons support the hypothesis of significant ($P < 0.05$) differentiation except the combination marked with *

	SAL	LAW	DUR	NAM	MOZ	KOL	KK
SAL	-	0.138	0.134	0.169	0.166	0.105	0.269
LAW	0.112	-	0.208	0.205	0.151	0.270	0.207
DUR	0.189	0.171	-	0.206	0.206	0.162	0.305
NAM	0.184	0.197	0.263	-	0*	0.162	0.358
MOZ	0.135	0.128	0.181	0.123	-	0.157	0.351
KOL	0.206	0.182	0.270	0.204	0.137	-	0.293
KK	0.167	0.140	0.239	0.219	0.088	0.167	-

SAL - SA Landrace; LAW - Large White; DUR - Duroc; NAM – Namibia; MOZ – Mozambique; KOL – Kolbroek; KK - Kune-Kune.

Results from the assignment test are presented in Figures 1 and 2. The probability of $K = 1 - 7$, averaged over 13 runs for each value of K , is presented in Figure 1. A continued gradual increase in $\Pr(X|K)$ values was observed for increased values of K , with the highest probability of K being 7. This interpretation was based on the combination of the highest $-\ln Pr$ value and a low SD, following Pritchard *et al.* (2000). The assignment of individuals to each cluster based on $K = 7$, from 1,000,000 MCMC iterations, is shown in Table 5. The proportion of membership to each cluster for $K = 1 - 7$ is presented in Figure 2.

Discussion

Differences among the levels of genetic diversity in populations correlated well with known population histories. In the commercial populations, the lowest Hz value (0.531) was found in the Duroc breed that was introduced into South Africa only 30 years ago and is still kept only by a limited number of farmers. Higher Hz values were found in the SA Landrace (0.580) that has been in the country since 1952 and is the second largest commercial breed; and in the Large White (0.636) which is the most prominent

local commercial breed and has been farmed since 1903. The Hz values in the latter two populations are comparable with values of 0.560 - 0.667 characterized as high by Martínez *et al.* (2000), Fan *et al.* (2002), Kim *et al.* (2005), SanCristobal *et al.* (2006), Thuy *et al.* (2006), Chang *et al.* (2009) and Lemus-Flores (2009) in European and Asian pig populations.

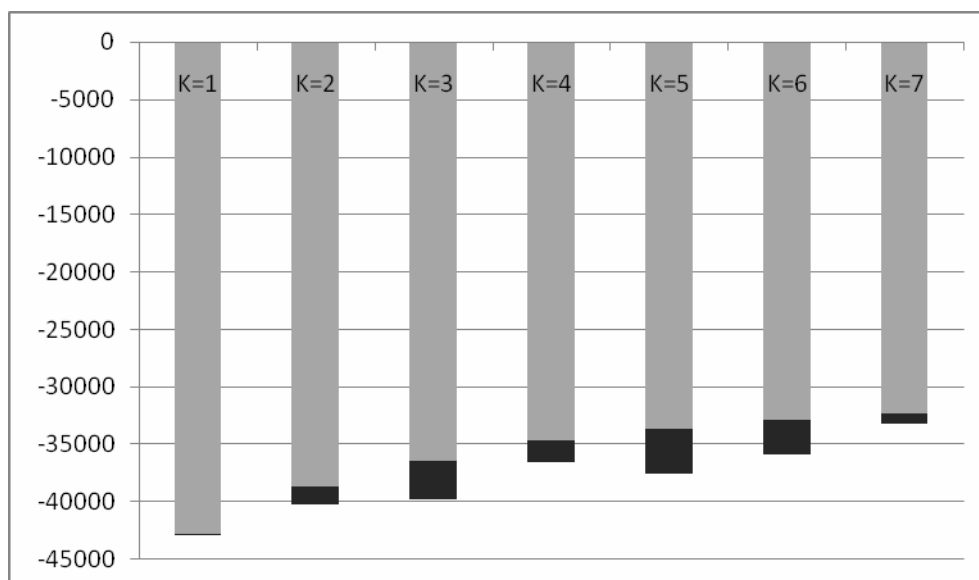


Figure 1 Probability (-LnPr) of K = 1 - 7, averaged over 13 runs (light grey bars), with standard deviation over 13 runs for each value of K (dark grey bars). For clarity of interpretation, standard deviation is multiplied by 10 in all cases.

Table 5 Proportion of membership of each pre-defined pig population in each of seven genetic clusters identified from Bayesian analysis

	Cluster:						
	1	2	3	4	5	6	7
SAL	0.003	0.014	0.004	0.958	0.008	0.009	0.004
LAW	0.002	0.008	0.004	0.966	0.013	0.003	0.004
DUR	0.002	0.002	0.002	0.017	0.974	0.002	0.001
NAM	0.002	0.978	0.001	0.011	0.001	0.002	0.005
MOZ	0.344	0.149	0.011	0.028	0.004	0.006	0.458
KOL	0.007	0.012	0.136	0.006	0.005	0.821	0.014
KK	0.023	0.027	0.569	0.323	0.013	0.003	0.042

SAL - SA Landrace; LAW - Large White; DUR - Duroc; NAM - Namibia; MOZ - Mozambique; KOL - Kolbroek; KK - Kune-Kune.

The trends from Hz values in commercial populations is confirmed by the average number of alleles per locus. The SA Landrace and Large White breeds showed A values of 5.475 - 5.5, which is comparable to values of 5.3 - 5.34 reported by Behl *et al.* (2006) and Rodríguez Rodríguez *et al.* (2008). The A values in the two South African commercial breeds is also higher than the average values of 4.59 reported for 11 European pig breeds by Laval *et al.* (2000) and 4.3 for Iberian pig breeds (Martinez *et al.*, 2000). The relatively high number of alleles found in the South African populations is an indication that the effects of

isolation and selection of these populations has been mild. The lower number of alleles in the Duroc population (3.975) supports the H_z value and reflects a relatively recently founded national population of limited size. The number of alleles in the South African Duroc pigs is, however, still higher than values of 2.39 - 2.8 reported in Belgian and some Asian populations (Van Zeveren *et al.*, 1995; Fan *et al.*, 2002; Chang *et al.*, 2009).

Historical information on origins, breeding and selection strategies applied to the indigenous pig populations of southern Africa is very limited. During the current study, the highest H_z levels were found in the Mozambican and South African (Kolbroek) populations, with values of 0.692 and 0.634, respectively. By comparison the value in the Namibian pigs was lower at 0.531. In the absence of reliable historical data it is not possible to fully evaluate these differences, but the trend does correlate with the fact that levels of pig farming in Namibia is low compared to South Africa and Mozambique. The authors note, however, that even the lowest H_z level of 0.531 in Namibia is comparable to the lowest value observed in a South African commercial breed (Duroc). The favourable values in the Mozambican and Kolbroek indigenous populations could reflect low selection pressure, natural patterns of mating, the presence of genetic substructure due to the many small farms sampled or a diverse founding history.

The average number of alleles in indigenous populations supports the trends from H_z . Values of 6.18 and 8.45 in the Kolbroek and Mozambican populations are higher than the highest value found in a commercial South African population (4.593). Since the high values in the Mozambican and Kolbroek populations may be due to bigger sample sizes, allelic richness was added as an additional measure for comparison. Using a calculated standardized sample size of 13 animals per population, R_s values in these two populations were 4.80 and 5.46. These values are still higher than the R_s values of 4.59 and 4.64 recorded in the two most diverse South African commercial populations, which confirm favourable levels of genetic diversity in these indigenous populations. In the Namibian group, the number of alleles ($A = 3.95$; $R_s = 3.58$) is significantly lower than that of the remaining indigenous populations but on par with levels of A and R_s in the Duroc commercial population.

Overall, the southern African commercial breeds and indigenous populations sampled, showed moderate to high levels of genetic diversity, as is also evident from the high number of private alleles found. This diversity should provide a valuable reservoir for both short-term management (parentage studies, individual identification and forensic applications) and future selection programmes.

The Kune-Kune population displayed levels of diversity ($H_z = 0.675$ and $A = 5.970$) that are comparable to the highest values obtained for southern African commercial and indigenous breeds. These high values can probably be contributed to low selection pressures. The Kune-Kune population also displayed a large number of private alleles. This population represents a unique reservoir of genetic diversity and appropriate conservation efforts should be applied.

The hierarchical division of total genetic diversity indicated that 17.9% of the total genetic variation was due to differences between populations. This high between-population component suggests high reproductive isolation and drift between breeds. The significance values of F_{ST} - and R_{ST} -based analysis provided additional confirmation of high levels of drift, with most pairwise comparisons supporting an hypothesis of significant ($P < 0.05$) differentiation.

Among the commercial pig populations, the highest F_{ST} values were found between the SA Landrace and the Duroc populations (0.189) and between the Large White and the Duroc populations (0.171). The F_{ST} value between the SA Landrace and the Large White populations was smaller at 0.112. This pattern supports results by Kotze & Visser (1996), who reported that SA Landrace pigs were more closely related to Large White and distant from the Duroc breed, using protein markers. This trend was anticipated because the Landrace breed was originally a cross between a native Danish pig and the Large White, whereas the Duroc represents a distinct breed. Results by Paszek *et al.* (1998) and (Lemus-Flores *et al.*, 2001), using linked microsatellites, also showed that Duroc animals were distant from other commercial populations, including the Large White. Among the indigenous populations, F_{ST} and R_{ST} values showed that the Kolbroek was slightly more differentiated from the Namibian and Mozambican populations, compared to differentiation between the latter two populations. In the absence of historical data, the mechanism for this cannot be determined.

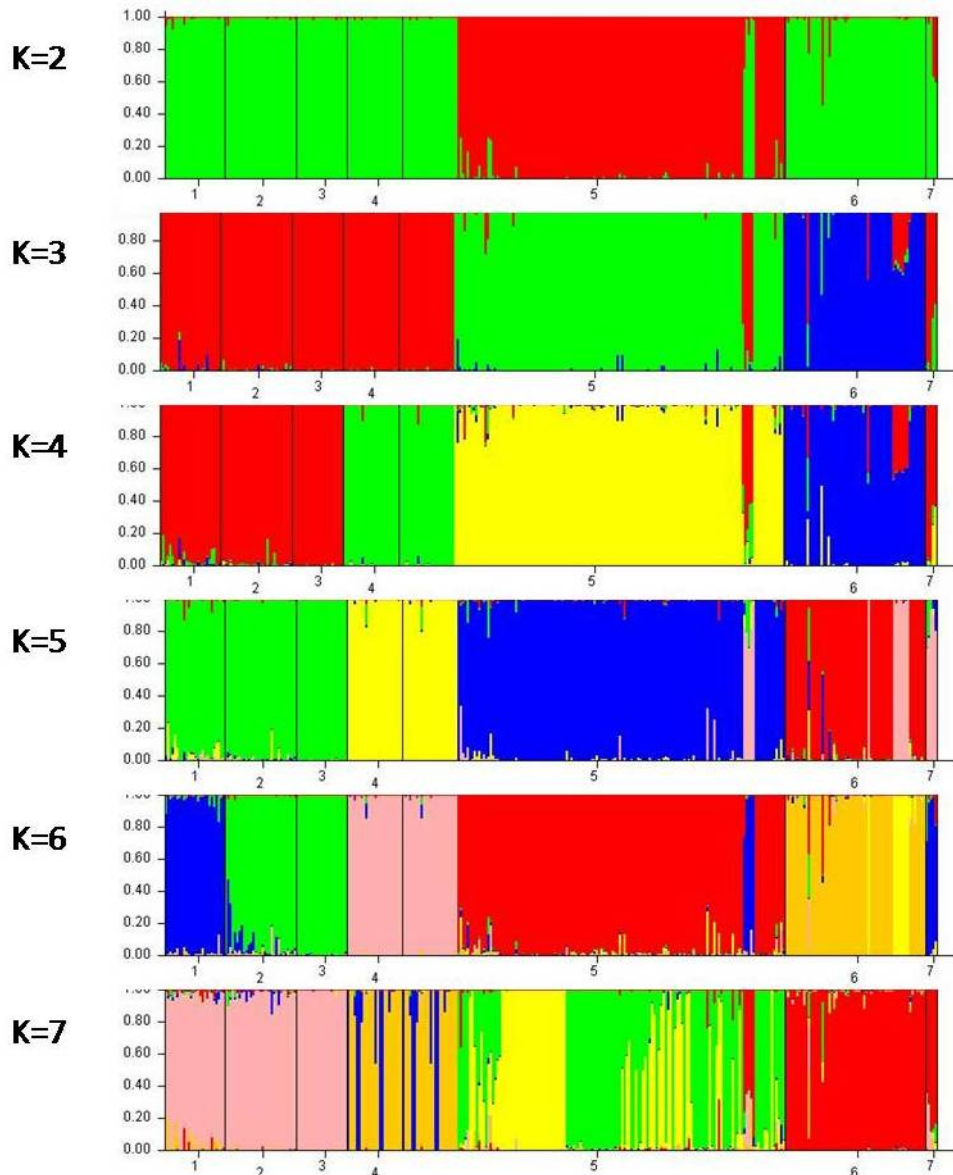


Figure 2 Proportion of membership of pig populations to 2 - 7 clusters, for possible real population structures consisting of 2 - 7 clusters. Order of populations on horizontal axis: 1 = SAL, 2 = LAW, 3 = DUR, 4 = NAM, 5 = MOZ, 6 = KOL and 7 = KK. (SAL - SA Landrace; LAW - Large White; DUR - Duroc (DUR); NAM – Namibia; MOZ – Mozambique; KOL – Kolbroek; KK - Kune-Kune).

Average F_{ST} values among commercial populations (0.157) and indigenous populations (0.155) were lower than values for pair-wise comparison of commercial and indigenous populations (0.194). However, pair-wise comparison showed unexpected low F_{ST} values between the Mozambican and Large White populations (0.128) and between the Mozambican and SA Landrace populations (0.135). Results from R_{ST} were generally similar. The trend of lower than expected differentiation between certain commercial and indigenous populations cannot be explained in the absence of historical data.

The F_{ST} values did not show a consistent pattern of high or low differentiation between the Kune-Kune breed and southern African populations (0.088 - 0.239). Trends from R_{ST} suggested comparatively more differentiation between the Kune-Kune and southern African populations ($R_{ST} = 0.207 - 0.358$ compared to values of 0 - 0.270 within and between commercial and indigenous groups).

Results from the assignment test suggested a true genetic structure with significant differentiation among all populations, except the commercial SA Landrace and Large White breeds. This result from the Bayesian-based assignment test therefore supports the trend from frequency-based F_{ST} values, which showed significant differentiation among all population pairs but with the lowest F_{ST} observed between the SA Landrace and Large White breeds. Five populations displayed a pattern of strong distinctiveness, with more than 80% of individuals assigned to a single cluster in each case. These were the SA Landrace (95.8%), Large White (96.6%), Duroc (97.4%), Namibia (97.8%) and Kolbroek (82.1%). By contrast, significant numbers of pigs from the Kune-Kune breed were distributed over several clusters, with the most prominent cluster containing only 56.9% of individuals. The Mozambican population showed less homogeneity still, with the most prominent cluster containing only 45.8% of pigs from that population. Overall, the results from the assignment test support the hypothesis of high homogeneity within most breeds or national populations sampled. The notable exception is the indigenous pigs from Mozambique. The distribution of pigs from this country over several clusters could reflect the large number of sub-populations sampled, or be the signature of a diverse founding history. In the absence of historical data, the exact mechanism cannot be determined.

Conclusions

This study represents the first attempt to genetically characterize domestic and indigenous pig populations in southern Africa using microsatellite markers. The results show that levels of genetic diversity in national populations of specific groups are moderate to high. This should provide a resource for selection and adaptation. Levels of diversity in individual populations were not assessed and breeders should be encouraged to practise sound management strategies. Estimates of differentiation and genetic structure confirm the history of individual populations where known, and suggest considerable uniqueness in indigenous pig populations. The results obtained from this study can now contribute to the establishment of routine DNA typing services to the advantage of the pig industry in South Africa. The practical importance of the results from the current study is, however, slightly compromised by the fact that the FAO-ISAG recommended panel of microsatellites was not used. It is thus recommended that further genetic characterization studies on southern African pigs should be undertaken using the FAO-ISAG recommended panel of microsatellites, to ensure compatibility between laboratories worldwide.

Acknowledgements

The authors wish to thank D. Visser for samples donated and for information on populations; and M. Rothschild from the Pig Genome Project, USA for donating the pig markers to the Animal Genetics Section at Irene. We thank the two anonymous reviewers for valuable comments that resulted in significant improvements to this MS.

References

- Alexander, L.J., Rohrer, G.A. & Beattie, C.R., 1996. Cloning and characterisation of 414 polymorphic porcine microsatellites. *Anim. Genet.* 27, 137-148.
- Behl, R., Sheoran, N., Behl, J. & Vijh, R.K., 2006. Genetic analysis of Ankamali pigs of India using microsatellite markers and their comparison with other domesticated Indian pig types. *J. Anim. Breed. Genet.* 123, 131-135.
- Bester, J. & Küsel, U.S., 1998. Early domesticated animals in South Africa. *Proc. 4th Global Conference on Conservation of Domestic Animal Genetic Resources, Nepal.*
- Briggs, H.M., 1983. *International Pig Breed Encyclopedia.* Elanco Animal Health, Indianapolis, USA.
- Budowle, B., Monson, K.L. & Chakraborty, R., 1996. Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. *Int. J. Legal Med.* 108, 173-176.
- Chang, W.H., Chu, H.P., Jiang, Y.N., Li, S.H., Wang, Y., Chen, C.H., Chen, K.J., Lin, C.Y. & Ju, Y.T., 2009. Genetic variation and phylogenetics of Lanyu and exotic pig breeds in Taiwan analyzed by nineteen microsatellite markers. *J. Anim. Sci.* 87, 1-8.
- DAS, 2010. Abstract of Agricultural Statistics. Compiled by the Directorate of Agricultural Statistics, Department of Agriculture, Forestry and Fisheries, South Africa. <http://www.daff.gov.za/>

- Ellegren, H., Chowdhary, B.P., Johansson, M., Marklund, L., Fredholm, M., Gustavsson, I. & Andersson, L., 1994. A primary linkage map of the porcine genome reveal low rate of recombination. *Genetics* 137, 1089-1100.
- El Mousadik, A. & Petit, R.J., 1996. High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. *Theor. Appl. Genet.* 92, 832-839.
- Epstein, H. & Masen, I.L., 1971. *The Origin of the Domestic Animals of Africa, Volume II.* Africana Publishing Corporations, New York.
- Excoffier, L., Laval, G. & Schneider, S., 2005. Arlequin, version 3.0: An Integrated Software Package for Population Genetics Data Analysis. Computational and Molecular Population Genetics Laboratory (CMPG), Institute of Zoology, University of Berne, Switzerland.
- Falush, D., Stephens, M. & Pritchard, J.K., 2003. Inference of population structure: extensions to linked loci and correlated allele frequencies. *Genetics* 164, 1567-1587.
- Fan, B., Wang, Z-G., Li, Y-J., Zhao, X-L., Liu, B., Zhao, S-H., Yu, M., Li, M-H., Chen, S-L., Xiong, T-A. & Li, K., 2002. Genetic variation analysis within and among Chinese indigenous swine populations using microsatellite markers. *Anim. Genet.* 33, 422-427.
- Fredholm, M., Wintero, A., Christensen, K., Kristensen, B., Brauer, N.P., Davies, W. & Archibald, A., 1993. Characterization of 24 porcine (dA-dC)n-9dT-dG)n microsatellites: Genotyping of unrelated animals from four breeds and linkage studies. *Mamm. Genome* 4, 187-192.
- GENESCAN[®] analysis software. Applied Biosystems. www.appliedbiosystems.com
- GENOTYPER[®] analysis software. Applied Biosystems. www.appliedbiosystems.com
- Goodman, S.J., 1997. RSTCALC: A collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Mol. Ecol.* 6, 881-885.
- Goudet, J., 2001. FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices. Version 293. Institut de Zoologie et d'Ecologie Animale, Universite de Lausanne, Lucerne, France.
- Groenen, M.A.M., Ruyter, D., Verstege, E.J.M., De Vries, M. & Van der Poel, J.J., 1995. Development and mapping of ten porcine microsatellite markers. *Anim. Genet.* 26, 115-118.
- Higuchi, R., Von Beroldingen, C.H., Sensabaugh, G.F. & Erlich, H.A., 1988. DNA typing from single hair. *Nature* 332, 543-546.
- Hoyheim, B., Keiserud, A. & Thomsen, P.D., 1994. A highly polymorphic porcine dinucleotide repeat S0301 (BHT 12) at chromosome 4p15. *Anim. Genet.* 25, 432.
- Kim, T.H., Kim, K.S., Choi, B.H., Yoon, D.H., Jang, G.W., Lee, K.T., Chung, H.Y., Lee, H.Y., Park, H.S. & Lee, J.W., 2005. Genetic structure of pig breeds from Korea and China using microsatellite loci analysis. *J. Anim. Sci.* 83, 2255-2263.
- Kotze, A. & Visser, D.P., 1996. Status of genetic variation in purebred pig breeds in South Africa. *Proc. All Afr. Conf. Anim. Agriculture*, Pretoria, 1-4 July.
- Lemus-Flores, C., Ulloa-Arvizu, R., Ramos-Kuri, M., Estrada, F.J. & Alonso, R.A., 2009. Genetic analysis of Mexican hairless pig populations. *J. Anim. Sci.* 79, 3021-3026.
- Lopez-Corrales, N.L., Beattie, C.W. & Rohrer, G.A., 1999. Cytogenetic assignment of 53 microsatellites from the USDA-MARC porcine genetic map. *Cytogen. Cell Genet.* 84, 140-144.
- Martínez, A.M., Delgado, J.V., Rodero, A. & Vega-Pla, J.L., 2000. Genetic structure of the Iberian pig breed using microsatellites. *Anim. Genet.* 31 (5), 295-301.
- McQueen, H.A., Yerle, M. & Archibald, A.L., 1994. Anchorage of an unassigned linkage group to pig chromosome 10 with p1 clones. *Mamm. Genome* 5, 646-648.
- Nei, M., 1987. *Molecular Evolutionary Genetics.* Second ed., Columbia University Press, New York, USA.
- Park, S., 2001. Microsatellite Toolkit. <http://oscar.gen.tcd.ie/sdepark/ms-toolkit>.
- Paszek, A.A., Flickinger, G.H., Fontanesi, L., Beattie, C.W. & Rohrer, G.A., 1998. Evaluating evolutionary divergence with microsatellites. *J. Mol. Evol.* 46, 121-126.
- Petit, R.J., El Mousadik, A. & Pons, O., 1998. Identifying populations for conservation on the basis of genetic markers. *Conserv. Biol.* 12, 844-855.
- Plug, I., 1993. Origin of South African livestock and their potential role in the industry. In: *Livestock Production Systems: Principles and Practice*, Eds Maree, C. & Casey, N.H., AGRI Development Foundation. pp. 11-13.

- Plug, I. & Badenhorst, S., 2001. The distribution of macromammals in Southern Africa over the past 30,000 years. Transvaal Museum Monograph 113, South Africa.
- Pritchard, J.K., Stephens, M., Rosenberg, N.A. & Donnelly, P., 2000. Association mapping in structured populations. *Am. J. Hum. Genet.* 67, 170-178.
- Prolit, 2004. Pig Farming. Project Literacy Productions, Cape Town, South Africa.
- Quin, P.J., 1959. Foods and Feeding Habits of the Pedi. Witwatersrand University Press, Johannesburg, South Africa.
- Ramsay, K.A., Reed, D.S., Bothma, A.J. & Lepen, J.M., 1994. Profitable and environmentally effective farming with early domesticated livestock in Southern Africa, Department of Agriculture, Pretoria.
- Rice, W.R., 1989. Analyzing tables of statistical tests. *Evolution* 43, 223-225.
- Robic, A., Dalens, M., Woloszyn, N., Milan, D., Riquet, J. & Gellin, J., 1994. Isolation of 28 new porcine microsatellites revealing polymorphism. *Mamm. Genome* 5, 580-583.
- Rodrigáñez, J., Barragán, C., Alves, E., Gortázar, C., Toro, M.A. & Silió, L., 2008. Genetic diversity and allelic richness in Spanish wild and domestic pig population estimated from microsatellite markers. *Span. J. Agric. Res.* 6, 107-115.
- Rohrer, G.A., Alexander, L.J., Keele, J.W., Smith, T.P.L. & Beattie, C.W., 1994. A microsatellite linkage map of the porcine genome. *Genetics* 136, 231-245.
- Rohrer, G.A., Alexander, L.J., Hu, Z., Smith, T.P.L., Keele, J.W. & Beattie, C.W., 1996. A comprehensive map of the porcine genome. *Genom. Res.* 6, 371-391.
- Rothschild, M.F. & Ruvinsky, A., 1998. *The Genetics of the Pig*. CABI Publishing, Oxfordshire, UK.
- Ruvinsky, A. & Rothschild, M.F., 1998. Systematics and phylogeny of pigs. In: *The Genetics of the Pig*. Eds Rothschild, M.F. & Ruvinsky, A., CAB International, UK. pp. 1-16.
- SanCristobal, M., Chevalet, C., Haley, C.S., Joosten, R., Rattink, A.P., Harlizius, B., Groenen, M.A.M., Amigues, Y., Boscher, M-Y., Russel, G., Law, A., Davoli, R., Russo, V., Désautés, C., Alderson, L., Fimland, E., Bagga, M., Delgado, J.V., Vega-Pla, J.L., Martiniz, A.M., Ramos, M., Glodek, P., Meyer, J.N., Gandini, G.C., Matassino, D., Plastow, G.S., Siggins, K.W., Laval, G., Archibald, A.L., Milan, D., Hammond, K. & Cardellino, R., 2006. Genetic diversity within and between European pig breeds using microsatellite markers. *Anim. Genet.* 37, 189-198.
- Thuy, N.T., Melchinger-Wild, E., Kuss, A.W., Cuong, N.V., Bartenschlager, H. & Geldermann, H., 2006. Comparison of Vietnamese and European pig breeds using microsatellites. *J. Anim. Sci.* 84, 2601-2608.
- Toro, M.A., Fernández, J. & Caballero, A., 2008. Molecular characterization of breeds and its use in conservation. *Anim. Genet. Res.* 120 (3), 174-195.
- Treacy, D.A., 1976. A genetic analysis of the pedigree Landrace pig breed in Australia. *Aust. J. Exp. Agric. Anim. Husband.* 16, 76-81.
- Van Zeveren, A., Peelman, L., Van de Weghe, A. & Bouquet, Y., 1995. A genetic study of four Belgian pig populations by means of seven microsatellite loci. *J. Anim. Breed. Genet.* 112, 191-204.
- Visser, D.P., 2004. Structuring of breeding objectives in the pork supply chain in South Africa. PhD thesis, University of Pretoria, South Africa.
- Visser, D.P., Browne, A. & Botes, F.DuT., 1993. Breeds and breeding programs. In: *Pig Production in South Africa*. Ed. Kemm, E.H., Agricultural Research Council. V & R Printers, Pretoria. pp. 19-20.