J. Anim. Breed. Genet. 118 (2001), 213–233
© 2001 Blackwell Wissenschafts-Verlag, Berlin
ISSN 0931–2668

Ms. received: 02.01.2001

[1]*School of Rural Science and Natural Resources, University of New England, Armidale, NSW 2351, Australia,* [2]*Department of Biology, Universiti Putra Malaysia, 43400 UPM, Serdang, Malaysia,* [3]*CSIRO Division of Tropical Animal Production, Molecular Animal Genetics Centre, University of Queensland, St Lucia, Queensland 4072, Australia and* [4]*Institute of Advanced Studies, University of Malaya, 59100 Kuala Lumpur, Malaysia*

Genetic variation within and relationships among populations of Asian goats (*Capra hircus*)

By J. S. F. Barker[1], S. G. Tan[2], S. S. Moore[3], T. K. Mukherjee[4], J.-L. Matheson[3] and O. S. Selvaraj[4]

Summary

Genetic variation at 59 protein coding loci (16 polymorphic) and 25 microsatellite loci was analysed for 11 indigenous south-east Asian goat populations, and the Australian feral population, to determine the magnitude of genetic differentiation and the genetic relationships among the populations. Significant deviations from Hardy–Weinberg equilibrium were detected in one or more populations for eight of the nine protein loci with codominant alleles, and for microsatellites for all except the two Sri Lankan populations and for all but four loci. For both marker types, average inbreeding coefficients (F_{IS}) were exceptionally high. Heterogeneity of deviations from Hardy–Weinberg equilibrium for the microsatellites showed no differences for among loci within populations as compared with among populations within loci. For protein loci, however, the former was higher, indicating selection affecting allele frequencies at some loci. The variance among protein loci was significantly higher than among microsatellite loci, further indicating selection at some protein loci. There was significant differentiation among populations for both protein and microsatellite loci, most likely reflecting the geography of south-east Asia, and the presumed spread of goats throughout the region. Phylogenies derived from pair-wise genetic distance estimates show some similar clustering for the microsatellite and protein based trees, but bootstrap support was generally low for both. A phylogeny based on the combined set of 38 protein and microsatellite loci showed better consistency with geography and higher bootstrap values. The genetic distance phylogeny and the Weitzman diversity tree derived from microsatellite data showed some identical clusters, and both identified the Ujung Pandang and Australia populations as contributing most to overall genetic diversity.

Keywords: Biochemical polymorphism, microsatellites, genetic distance, phylogeny, inbreeding

Introduction

The indigenous goats of south-east Asia are not classified into breeds, and all are similar phenotypically to the Kambing Katjang of Indonesia and Malaysia (Devendra and Nozawa 1976). However, since their spread through the region, local populations may have been largely isolated, and thus may now be genetically differentiated. The majority of the goats are kept by small holder farmers, primarily for meat production, and mating is uncontrolled and indiscriminate (Devendra and Nozawa 1976). Although generally secondary to other farm enterprises, goat production may contribute substantially to farm income, particularly for small holders. Nevertheless, the breeding and improvement of indigenous goats has received relatively little attention in the region (Food and Fertilizer Technology Center 1984), with only limited study of crossbreeding with exotic breeds (Panandam et al. 1987; Hirooka et al. 1997).

There is a growing world-wide recognition of the need for conservation of livestock diversity (FAO 1995), and for the characterization of breeds and populations, including their genetic differentation and relationships. Although there have been some studies of blood group and protein variation in Asian goats (NOZAWA et al. 1978; KATSUMATA et al. 1982), and of breed relationships using microsatellites (YANG et al. 1999), no systematic study has considered populations across the region. We present here an analysis of genetic structure and relationships of 12 geographic populations, using both protein coding and microsatellite loci.

Materials and methods

Experimental animals

The populations (Fig. 1) and numbers of animals studied for protein variation (600 in total – 80 males, 520 females) and for microsatellite variation (a subset of 308) were: Sri Lanka – Sri Lanka South (52, 26), Sri Lanka North-Central (31, 30); Thailand – Chiang Mai (50, 31), Hat Yai (39, 32); Malaysia – Malaysian Agriculture Research and Development Institute/ University of Malaya (55, 0 – samples lost before DNA extraction due to freezer breakdown), Sabah (51, 26), Sarawak (71, 24); Philippines – Musuan (51, 26); Indonesia – Medan (50, 26), Bogor (50, 25), Ujung Pandang (48, 27); Australia (52, 35). The Australian population derived from some 1000 feral goats collected in western New South Wales in 1978 and 1981, and subsequently maintained as a closed flock. All other populations were indigenous goats. With the exception of Medan, Hat Yai, Australia and Sri Lanka North-Central, animals were sampled from two or more subpopulations, generally different villages in the same area (except Hat Yai and Malaysia – institutional flocks). In all cases, owners were questioned in detail to minimize the sampling of closely related individuals.

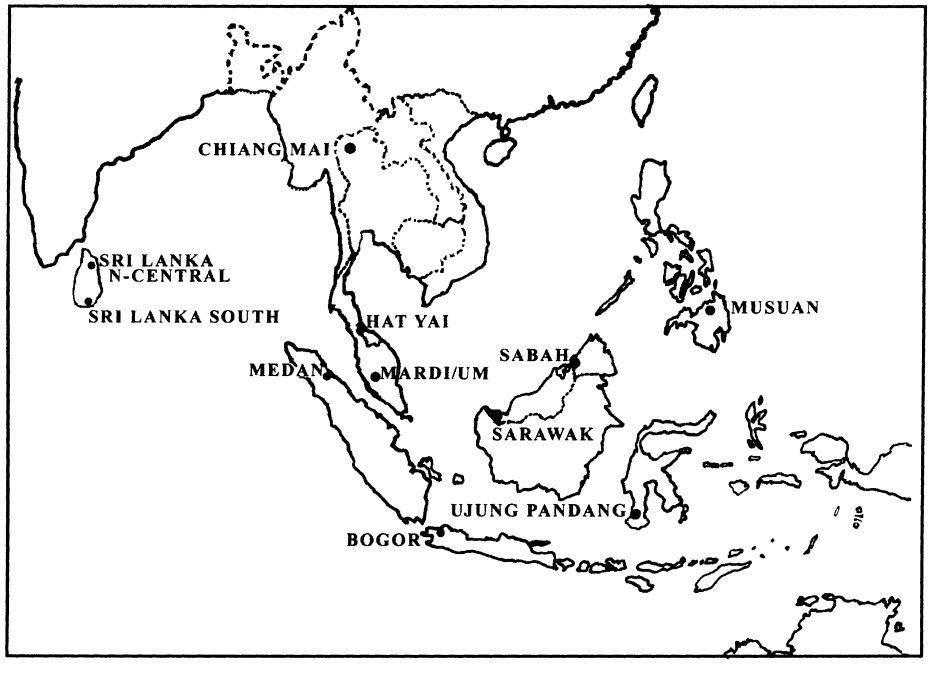

Fig. 1. Map of South-East Asia, with locations shown for the 12 populations sampled

Twenty millilitres of whole blood was collected from each animal into vacuum tubes with heparin as anticoagulant. The blood was centrifuged for 5 min and plasma, white cell layer (buffy coat) and red cells were suctioned separately into 2 ml vials. These vials were immediately frozen in liquid nitrogen in the field, and transferred later to a –80°C freezer in the laboratory.

DNA extraction

DNA was extracted from frozen white blood cells. To each sample 2.5 volumes of 0.01 M KCl, 0.15 M NH_4Cl, 0.1 mM EDTA was added and the sample kept on ice for 10 min. The nuclear pellet was spun down at 4000 rpm at 4°C and washed twice with 10 ml of Tris buffered saline (0.14 M NaCl, 5 mM KCl, 0.25 mM Tris pH 7.4). The pellet was suspended in 9 ml of TE buffer and then 500 μl of 0.5 M EDTA (pH 8), 40 μl of Proteinase-K (20 mg/ml) and 500 μl of 10% SDS was added. The resulting nuclear lysate was incubated at 55°C for 3–4 h or overnight at 37°C. Following incubation 3 ml of 5 M NaCl and 13 ml of chloroform was added. The sample was mixed slowly on a rotator for 30 min and then the precipatated proteins spun down at 4000 rpm for 30 min at room temperature. The top aqueous layer was recovered, being careful not to disturb the protein layer at the chloroform interface. DNA was precipitated by gentle addition of 2.5 volumes of ethanol. The resulting DNA strands were spooled out and washed twice with ice cold 70% ethanol to remove excess salt. DNA was redissolved in 1–2 ml of TE buffer pH 8.0.

Protein loci

The 59 loci analysed and electrophoretic assay systems are given in Appendix 1. Assay and staining methods for most are given by TAN et al. (1990, 1991, 1993) and BARKER et al. (1997a). Electrophoretic assay systems for other loci were: Blvr – 0.1 M Phosphate pH 7.0 buffer, 100 V for 2 h at 4°C, stain (RICHARDSON et al. 1986); Dpgm – 0.1 M Tris, 0.1 M NaH_2PO_4, pH 7.4 buffer, 200 V for 1 h, stain (HARRIS and HOPKINSON 1976); Gsr – 0.13 M Tris–EDTA-borate pH 8.9 buffer, 200 V for 2 h, stain (RICHARDSON et al. 1986); Sordh – 0.1 M Tris–HCl pH 8.0 buffer, 200 V for 100 min, stain (RICHARDSON et al. 1986). Known standards were included on all gels to ensure consistency of genotype scoring. When null phenotypes were detected, the assay was repeated to confirm absence of activity.

For each locus, 25 animals from each population were assayed initially, and 43 loci were found monomorphic in all populations (Appendix 1). Null alleles were detected at five loci (c Amy-1, Ca, Dia_1–1, Dia_1–2, and Xp), and variants at Amy and Np showed high and low activity, with high dominant. Mendelian inheritance with codominant alleles has been demonstrated from family studies for Alb, Hb, Tf and Xp (see TAN et al. 1991), and for Ca, Dia_1–1, Dia_1–2, Mdh and Me (unpublished data).

Microsatellite loci

The 25 microsatellite markers used (Appendix 2) were selected from a panel of 51 on the basis of amplification, and ability to coload based on product size and dye label. Three (MAF70, OarFCB48 and OarFCB193) are ovine markers, the rest are bovine, and all are included in the linkage map analysis of VAIMAN et al. (1996). Genotyping was carried out using an Applied Biosystems 373 DNA sequencer. One of each PCR primer pair was labelled with proprietry dye TET, 6FAM or HEX (Perkin Elmer/ABI). PCR reactions were carried out as previously described (MOORE et al. 1994) with some adjustment of annealing temperature in some cases to produce strong and specific PCR product in goats (Appendix 2). PCR products were coloaded as indicated in Appendix 2 and separated on a 12-cm well to read plates. Data was captured using GeneScan software (Perkin Elmer/

ABI). Data analysis was carried out using Genotyper 2.0 software (Perkin Elmer/ABI) and independently checked manually for genotyping errors. Genotypic data was then assembled into Excel spreadsheets for further analysis. Three loci (INRA005, INRA063, INRA071) were not assayed in one or more populations because of degraded or insufficient DNA.

Allelic frequency and heterozygosity

For the five protein loci where null homozygotes were detected in one or more populations, genotype and allele frequencies were estimated using the computer program GENEPOP Version 3.1d (RAYMOND and ROUSSET 1995) separately for each subpopulation where null homozygotes were detected, and weighted (by number of animals in each subpopulation) overall allele frequencies estimated. Allele frequencies at the two loci where one allele is dominant were computed using the bias correction of LYNCH and MILLIGAN (1994). Thus for these seven loci, Hardy–Weinberg equilibrium is assumed.

For the nine protein loci where the genotypes of each animal were determined directly (i.e. excluding loci with null alleles, or dominance) and for the 22 microsatellite loci assayed in all populations, observed and expected heterozygosity estimates and average inbreeding coefficients for each population were obtained using the BIOSYS-1 computer program (SWOFFORD and SELANDER 1989), but tests for deviations from Hardy–Weinberg equilibrium were done using the exact tests of GENEPOP. Significance levels for each test were determined by applying to the probability estimates calculated by GENEPOP the sequential Bonferroni procedure (HOCHBERG 1988; LESSIOS 1992) over loci within each population. Heterogeneity of deviations from Hardy–Weinberg equilibrium for the protein loci was examined among loci within each population and among populations for each locus by treating the deviations as correlation coefficients and testing accordingly (GAFFNEY et al. 1990; SOKAL and ROHLF 1981).

Statistical analyses

For those populations where samples were obtained from two or more subpopulations, deviations from Hardy–Weinberg equilibrium in the population could be due to genetic differences between subpopulations, and a consequent Wahlund effect. Thus tests of Hardy–Weinberg equilibrium and genotypic differentiation between subpopulations were done using GENEPOP. For those population/locus combinations where genotypic differentiation was not significant, overall allele frequencies for the population were used in subsequent analyses. Where genotypic differentiation among subpopulations was significant, then (i) if Hardy–Weinberg was not significant for each subpopulation or for the population, overall allele frequencies for the population were used, or (ii) if Hardy–Weinberg was not significant for all or most subpopulations, but significant for the population, overall allele frequencies were used, but the Wahlund effect on genotype frequencies was computed. Given the observed allele frequencies in each subpopulation, the expected heterozygote deficit due to the Wahlund effect can be computed (LI 1976). Expressing this as a percentage of the observed heterozygote deficit then measures the contribution of the Wahlund effect to the observed heterozygote deficit.

Tests for conformance of the protein-coding loci to neutral expectations were done using the ARLEQUIN package (SCHNEIDER et al. 2000). Tests for pair-wise linkage (genotypic) disequilibria among the microsatellite loci, separately for each population and overall, were done using GENEPOP, and then applying the sequential Bonferroni procedure to determine significance levels. For the microsatellite loci and for the nine protein loci whose genotypes were determined directly, F-statistics and their significance were determined using FSTAT Version 2.7 (GOUDET 1995), with the sequential Bonferroni procedure applied over loci in deriving significance levels. These parameters of population structure

are defined as the correlations between pairs of genes (i) within individuals (F) (ii) between individuals in the same population (θ), and (iii) within individuals within populations (f), and are analogous to WRIGHT's (1951, 1978) F_{IT}, F_{ST} and F_{IS}, respectively. F-statistics for the seven protein loci with null alleles or dominance were computed using the methods of WEIR and COCKERHAM (1984), as implemented in the computer program DIPLOID (WEIR 1990), adapted for input of the numbers of each genotype at each locus. Hierarchical F-statistics were computed using the methods of WEIR and COCKERHAM (1984), as implemented in the computer program GDA (LEWIS and ZAYKIN 1999), with Musuan and Australia deleted as each had only one population sampled.

Mean gene diversity for each locus and genetic distances among populations (standard genetic distance of NEI (1978) and the D_A distance of NEI et al. (1983) were obtained using the DISPAN computer program (T. OTA, personal communication). REYNOLDS' distance (REYNOLDS et al. 1983) and the $(\delta\mu)^2$ distance (GOLDSTEIN et al. 1995) were obtained using the MICROSAT computer program (MINCH et al. 1995). In addition, the D_A microsatellite distances were used to measure diversity, as proposed by WEITZMAN (1992, 1993), and to evaluate expected losses of diversity due to extinction of individual populations (THAON D'ARNOLDI et al. 1998).

Isolation by distance was tested using data on the microsatellite loci and on the nine polymorphic protein loci where the genotypes of each animal were determined directly. Estimates of θ (WEIR and COCKERHAM 1984) for each pair of populations were computed, and pair-wise $\theta/(1 - \theta)$ values then were regressed on log (geographic distance) between each pair of populations (ROUSSET 1997), and the significance of the association estimated using MANTEL's (1967) permutation test. A significant association indicates genetic structuring, and limited dispersal.

Results

For the 16 polymorphic protein loci, some populations show markedly different allele frequencies for some loci, although there are few cases of consistent patterns for different loci. One apparent pattern is for Australia and the Sri Lankan populations, where Hb^B and Tf^A have very high frequencies, whereas Hb^A shows highest frequency and Tf^A is at much lower frequency in all other populations. Eleven loci were polymorphic in all 12 populations, while one locus (Dia_1–1) was polymorphic in only five populations. The total number of alleles detected at the microsatellite loci ranged from five (ILSTS005) to 18 (TGLA179). Two loci (BL25 and ILSTS005) were each fixed in one population, while BM757 was fixed in six populations. As for the protein loci, some populations show markedly different microsatellite allele frequencies for some loci. In fact, for 19 of the 25 loci, the overall most frequent allele was not the most frequent in one or more populations, and in some cases, was not even present in one population. Across the 25 loci, 85 of the 295 alleles detected were unique to one population, with these alleles generally those at the lower or upper ends of the size distribution, and occurring only once. The Ujung Pandang population was the most extreme, with 28 alleles of 283 (this population was not assayed for INRA071) unique to this population, including three alleles of seven for BM757 and eight of 15 for TGLA272. Tables of allele frequencies for the polymorphic protein and the microsatellite loci and the primary data files (genotypes of all animals) are available at http://ansc.une.edu.au/ansc/genetics/– link through Barker, J. S. F. Measures of genetic variation for each population (mean number of alleles per locus and observed and expected heterozygosity) are given in Table 1. For all populations, average observed heterozygosities are less than the expected. Average gene diversity within populations (NEI 1973) for the protein loci ranged from 0.117 (Cat) to 0.572 (Xp), with an overall mean of 0.339, and for the microsatellite loci from 0.259 (INRA003) to 0.702 (MAF070), with an overall mean of 0.520.

Table 1. **Sample sizes and genetic variability in all populations (standard errors in parentheses)**

	Protein loci				Microsatellite loci (22 loci)			
	All 16 polymorphic loci		Mean heterozygosity (9 loci)[a]				Mean heterozygosity	
Population	Mean sample size per locus	Mean no. of alleles per locus	Direct count	Hardy–Weinberg expected[b]	Mean sample size per locus	Mean no. of alleles per locus	Direct count	Hardy–Weinberg expected[b]
Sri Lanka South	47.3 (0.43)	2.3 (0.21)	0.193 (0.042)	0.323 (0.063)	21.4 (0.8)	5.6 (0.5)	0.477 (0.060)	0.563 (0.055)
Sri Lanka N-Central	30.9 (0.06)	2.1 (0.14)	0.195 (0.044)	0.264 (0.048)	14.4 (1.0)	5.3 (0.5)	0.486 (0.053)	0.598 (0.052)
Chiang Mai	48.4 (0.68)	2.3 (0.17)	0.269 (0.068)	0.349 (0.054)	23.4 (0.8)	4.7 (0.5)	0.393 (0.042)	0.496 (0.043)
Hat Yai	38.6 (0.26)	2.2 (0.16)	0.342 (0.080)	0.383 (0.059)	23.4 (0.9)	4.5 (0.4)	0.427 (0.049)	0.494 (0.052)
MARDI/UM	52.6 (0.96)	2.4 (0.18)	0.175 (0.048)	0.228 (0.054)	–	–	–	–
Sabah	49.4 (0.68)	2.3 (0.17)	0.195 (0.056)	0.250 (0.055)	17.6 (1.0)	4.4 (0.4)	0.308 (0.045)	0.430 (0.043)
Sarawak	64.4 (2.54)	2.4 (0.18)	0.243 (0.066)	0.296 (0.066)	17.9 (1.2)	4.5 (0.4)	0.329 (0.048)	0.458 (0.052)
Musuan	47.5 (1.35)	2.5 (0.18)	0.339 (0.088)	0.382 (0.052)	21.6 (0.6)	5.3 (0.5)	0.447 (0.058)	0.576 (0.058)
Medan	47.1 (1.12)	2.4 (0.18)	0.254 (0.049)	0.336 (0.041)	16.8 (1.0)	4.3 (0.4)	0.417 (0.057)	0.478 (0.050)
Bogor	49.1 (0.45)	2.3 (0.15)	0.241 (0.073)	0.305 (0.058)	21.3 (0.7)	5.9 (0.7)	0.476 (0.047)	0.556 (0.050)
Ujung Pandang	47.0 (0.42)	2.4 (0.18)	0.218 (0.050)	0.355 (0.057)	20.9 (0.8)	5.6 (0.5)	0.379 (0.045)	0.581 (0.042)
Australia	51.8 (0.17)	2.3 (0.18)	0.185 (0.030)	0.315 (0.046)	19.5 (2.0)	5.4 (0.5)	0.451 (0.049)	0.662 (0.049)

[a]Loci whose genotypes were enumerated completely. [b]Unbiased estimate (see NEI 1978)

Significant deviations from Hardy–Weinberg equilibrium were detected in one or more populations for eight of the nine protein loci where genotypes were enumerated completely (Table 2). Three loci (Gc, Mdh, Me) showed significant deviations in eight or more populations. All populations had both positive and negative significant deviations, while deviations for each locus across populations were consistently positive (observed heterozygote deficiency) for Gc, Mdh and Me. For Alb, all populations showed negative deviations, except for Ujung Pandang (significantly positive). Observed heterozygote excesses (negative f) were found for Lap-1 in all but two populations. Tests for heterogeneity of deviations from Hardy–Weinberg equilibrium for the protein loci (Table 3) show greater heterogeneity among loci (higher χ^2 values) within each population than among populations for each locus. For the microsatellite loci, significant deviations from Hardy–Weinberg equilibrium were detected in all except the two Sri Lankan populations, and for all but four loci (Table 2). All significant deviations except one (TGLA325 in Medan) were positive (observed heterozygote deficiency). For both protein and microsatellite loci, Wahlund effects account for less than 40% of the observed heterozygote deficits for a few locus/population combinations (Table 2), and average inbreeding coefficients (F_{IS}) for each population are very high (Table 4). Tests for heterogeneity of deviations from Hardy–Weinberg equilibrium for the microsatellite loci had similar total χ^2 values for heterogeneity among loci and for among populations. For both protein and microsatellite loci, Ujung Pandang had the highest frequency of significant deviations. Strong evidence for selection affecting allele frequencies at protein-coding loci was found for Amy (one population, $p < 0.01$; eight populations, $0.05 < p < 0.10$), Np (seven – $p < 0.05$; two – $0.05 < p < 0.10$), and Xp (seven – $p < 0.05$; three – $0.05 < p < 0.10$). Significant effects ($p < 0.05$) in two – five populations were found for Ca, Dia_1–2, Lap-1 and Mdh. In all cases, the observed allele frequency distributions were more even than expected under the neutral model.

Significant linkage disequilibria ($p < 0.05$) were detected in the overall microsatellite data for 16 locus pairs (of a possible 300). For individual populations, the numbers of tests significant after Bonferroni correction were: Ujung Pandang – 7, Musuan – 5, Sri Lanka South – 2, and Sarawak – 1. Only one of the significant cases (in both overall and Ujung Pandang) was for markers on the same chromosome (TGLA272 and TGLA378 – unlinked on chromosome 10).

Overall means for the F-statistics are significantly different from zero, while those estimated for the nine protein loci with codominant alleles and for the microsatellite loci (Table 5) are not significantly different. For the seven protein loci with null alleles or dominance, Hardy–Weinberg was assumed in estimating genotype frequencies, so f estimates are not expected to differ from zero. The significant F-value for Dia_1–2, and high but not significant estimates for Ca and Dia_1–1 most likely result from undetected null heterozygotes in some populations.

To quantify genetic differentiation among populations for microsatellite loci, SLATKIN (1995) derived the R_{ST} statistic under the assumptions of the generalized stepwise mutation model. As we wish to compare the microsatellite results with those from the protein loci, and as R_{ST} and WEIR and COCKERHAM'S (1984) θ statistic are not expected to be greatly different for short-term differentiation of populations (SLATKIN 1995), we present results in terms of θ. As a check, however, both θ and ρ (an estimator of R_{ST}, GOODMAN 1997) were estimated for the 13 microsatellite loci where the total length of the flanking sequence was known. Over all loci, θ and ρ were similar (0.132 and 0.128, respectively), and for pair-wise estimates, the regression coefficient for ρ on θ was 0.960 ± 0.163. Analyses of population differentiation (θ in Table 5) show significant differentiation among populations for both protein and microsatellite loci, but with mean estimates for each class of locus not significantly different. As pointed out by ESTOUP and ANGERS (1998) and HEDRICK (1999), the high level of polymorphism for microsatellite loci is expected to increase the power of exact tests for population differentiation, and to decrease the values of F_{ST}, as compared

Table 2. **Significant tests of deviation from Hardy–Weinberg equilibrium. Figures given for significant tests are *f*-values (WEIR and COCKERHAM 1984)**

Locus	Sri Lanka South	Sri Lanka N-Central	Chiang Mai	Hat Yai	MARDI/ UM	Sabah	Sarawak	Musuan	Medan	Bogor	Ujung Pandang	Australia
Protein loci												
Alb											0.747	
Cat			A			A	A					
Gc	0.671		0.294[a]		0.408			0.325	0.518	0.410	0.578	0.386
Hb			0.561								0.461	
Mdh		1.000	0.511		0.758	0.831[a]	0.716[a]		0.499	0.822	0.409	0.868
Me	0.858	0.464	0.594	0.514		0.619	0.590	0.911	0.658	0.833	0.605	0.798
Lap-1								−0.550				
Tf	A										0.142	
Pt-Tf			0.507			A	A		0.441	A		
Microsatellite loci												
BL25				A	C			B			0.787	
BM757		A	A	A	C	B	A	A	B		0.519	A
BM3205					C					0.564	0.558	
BMC1009					C		0.660[a]					
CSSM026					C		0.398[a]	0.109				0.847
ILSTS005					C		A		B			
INRA003					C			B		B	0.726	0.524
INRA005					C		C				0.466	
INRA063		C			C						0.285	C
INRA071					C		C	0.813	C	C	C	C
INRA177					C							
MAF70					C							
OarFCB48					C						0.164	
OarFCB193					C		0.236		1.000		0.435	

Table 2. **continued**

Locus	Sri Lanka South	Sri Lanka N-Central	Chiang Mai	Hat Yai	MARDI/ UM	Sabah	Sarawak	Musuan	Medan	Bogor	Ujung Pandang	Australia
RM004					C			0.468		0.625	1.000	0.576
RM006					C	0.484[a]	0.709	0.509		0.190	0.415	0.700
RM044					C	0.546						
RM188					C							
RME23					C						0.211	
TGLA75			B		C		0.644	0.223			0.492	
TGLA179				0.598	C				0.592	0.370	0.352	
TGLA245					C	0.837	0.792	0.907	B			
TGLA272					C				0.283		0.654	
TGLA325			0.597	0.316	C	0.654			−0.827		0.178	
TGLA378					C		0.319	0.268			0.427	

A, monomorphic; B, only 1 heterozygote in sample; C, not assayed. [a]Wahlund accounts for 9–20% of observed heterozygote deficit for protein loci, and 30–40% for microsatellite loci

Table 3. **Tests for heterogeneity of deviations from Hardy–Weinberg equilibrium for the protein loci**

Heterogeneity among loci for each population			Heterogeneity among populations for each locus		
Population	d.f.	χ^2	Locus	d.f.	χ^2
Sri Lanka South	6	71.13***	Alb	10	58.90***
Sri Lanka N-Central	7	347.98***	Cat	8	21.33**
Chiang Mai	6	44.34***	Gc	10	26.11**
Hat Yai	7	38.21***	Hb	10	41.27***
MARDI/UM	7	53.86***	Mdh	10	310.33***
Sabah	5	80.89***	Me	10	60.13***
Sarawak	6	61.41***	Lap-1	10	37.43***
Musuan	7	148.48***	Tf	9	6.77
Medan	7	46.84***	Pt-Tf	8	29.00***
Bogor	7	122.07***			
Ujung Pandang	7	43.81***			
Australia	7	103.30***			

**p < 0.01, ***p < 0.001

Table 4. **Average inbreeding coefficients (f = F_{IS}) estimated for each population from data on nine protein loci and 22 microsatellite loci**

Population	Protein loci	Microsatellite loci
Sri Lanka South	0.397	0.127
Sri Lanka N-Central	0.252	0.150
Chiang Mai	0.226	0.190
Hat Yai	0.095	0.119
MARDI/UM	0.227	–
Sabah	0.212	0.257
Sarawak	0.183	0.249
Musuan	0.109	0.206
Medan	0.233	0.090
Bogor	0.201	0.118
Ujung Pandang	0.382	0.333
Australia	0.407	0.277

with allozyme loci. For these goat populations, there is little or no difference between the two classes of markers, either for F_{ST} or exact tests of population differentiation. All microsatellite and the nine protein loci with codominant alleles had θ-values significantly greater than zero (GENEPOP exact tests), and only two loci showed no significant differentiation among populations (Amy and Xp – *t*-test using jackknife estimates, Table 5). Further, all pair-wise comparisons (across all loci) of populations for genotypic frequencies were highly significant ($p < 0.001$) for microsatellites, and (with exceptions of Hat Yai/Musuan, Bogor/Ujung Pandang and Medan/Ujung Pandang – $p < 0.01$, Chiang Mai/Musuan – $p < 0.05$) for protein loci.

Population subdivision was analysed among countries (θ_P) and among populations within countries (θ_S). For the nine protein loci where genotypes were enumerated completely, estimates were 0.086 and 0.131, respectively, and for the 25 microsatellite loci 0.049 and 0.150, with all estimates significantly greater than zero.

The regressions of $\theta/(1 - \theta)$ on log (geographic distance), estimated separately for the protein and microsatellite loci, but excluding Australia as a nonindigenous population, were not significant, either for all pairs of populations, or when excluding those < 1500 km

Table 5. **F-statistics analyses for (1) nine allozyme loci (all genotypes defined (2) seven allozyme loci with null homozygotes or dominance and (3) 25 microsatellite loci. For (1) and (3), significance from permutation tests in the FSTAT program; for (2), significance from *t*-test using estimated standard deviation**

Locus[a]	f (F_{IS})	Θ (F_{ST})	F (F_{IT})
Protein loci			
Alb	−0.105 (0.105)	0.045 (0.020)***	−0.056 (0.099)
Cat	0.148 (0.063)*	0.069 (0.031)***	0.208 (0.074)***
Gc	0.369 (0.063)***	0.076 (0.049)***	0.415 (0.044)***
Hb	0.258 (0.073)***	0.449 (0.100)***	0.589 (0.074)***
Mdh	0.589 (0.067)***	0.036 (0.021)***	0.603 (0.062)***
Me	0.679 (0.045)***	0.069 (0.036)***	0.702 (0.047)***
Lap-1	−0.124 (0.083)	0.112 (0.043)***	0.001 (0.082)
Tf	−0.015 (0.042)	0.111 (0.047)***	0.097 (0.059)*
Pt-Tf	0.323 (0.082)***	0.035 (0.016)***	0.347 (0.079)***
Mean[b]	0.256 (0.120)***	0.121 (0.047)***	0.345 (0.105)***
Loci for which exact tests of significance are not possible (dominance or null alleles)			
Amy	0.009 (0.005)	0.009 (0.009)	0.017 (0.009)
Ca	0.179 (0.103)	0.149 (0.039)***	0.304 (0.109)**
c Amy-1	−0.013 (0.013)	0.154 (0.035)***	0.143 (0.034)***
Dia_1–1	0.159 (0.204)	0.225 (0.088)*	0.339 (0.140)*
Dia_1–2	0.217 (0.089)*	0.201 (0.074)**	0.377 (0.108)***
Np	−0.005 (0.004)	0.024 (0.011)*	0.019 (0.010)
Xp	0.058 (0.039)	0.095 (0.072)	0.149 (0.089)
Mean[b]	0.085 (0.039)*	0.106 (0.031)***	0.183 (0.062)**
Microsatellite loci			
BL25	0.329 (0.129)***	0.079 (0.027)***	0.384 (0.130)***
BM757	0.468 (0.315)**	0.109 (0.066)***	0.553 (0.356)***
BM3205	0.281 (0.071)***	0.131 (0.026)***	0.375 (0.066)***
BMC1009	0.155 (0.058)***	0.115 (0.039)***	0.253 (0.066)***
CSSM026	0.172 (0.077)***	0.178 (0.063)***	0.319 (0.077)***
ILSTS005	0.278 (0.108)***	0.085 (0.032)***	0.340 (0.101)***
INRA003	0.408 (0.126)***	0.442 (0.271)***	0.687 (0.229)***
INRA005	0.206 (0.062)***	0.093 (0.028)***	0.278 (0.046)***
INRA063	0.249 (0.069)***	0.208 (0.096)***	0.401 (0.062)***
INRA071	0.502 (0.109)***	0.207 (0.031)***	0.607 (0.096)***
INRA177	0.123 (0.039)**	0.116 (0.031)***	0.224 (0.044)***
MAF70	0.110 (0.039)**	0.068 (0.020)***	0.170 (0.040)***
OarFCB48	0.161 (0.058)***	0.175 (0.051)***	0.308 (0.070)***
OarFCB193	0.184 (0.066)***	0.123 (0.059)***	0.285 (0.086)***
RM004	0.429 (0.104)***	0.156 (0.109)***	0.519 (0.111)***
RM006	0.262 (0.079)***	0.108 (0.033)***	0.342 (0.071)***
RM044	0.132 (0.058)***	0.104 (0.023)***	0.223 (0.060)***
RM188	0.210 (0.051)***	0.179 (0.040)***	0.352 (0.058)***
RME23	0.143 (0.034)**	0.189 (0.050)***	0.305 (0.051)***
TGLA75	0.220 (0.065)***	0.125 (0.046)***	0.317 (0.058)***
TGLA179	0.344 (0.058)***	0.086 (0.024)***	0.400 (0.052)***
TGLA245	0.490 (0.174)***	0.130 (0.069)***	0.548 (0.133)***
TGLA272	0.146 (0.090)**	0.221 (0.088)***	0.340 (0.136)***
TGLA325	0.182 (0.113)***	0.143 (0.034)***	0.298 (0.093)***
TGLA378	0.150 (0.057)***	0.150 (0.030)***	0.277 (0.049)***
Mean[b]	0.230 (0.023)***	0.143 (0.010)***	0.340 (0.023)***

[a]Standard deviations in parentheses – estimate from jackknife over populations, [b]Standard deviations in parentheses – estimate from jackknife over loci. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

apart. Up to 1500 km, $\theta/(1-\theta)$ values were < 0.2, but for greater distances the values ranged from zero to 0.473 for the protein loci and 0.082–0.278 for the microsatellite loci.

TAKEZAKI and NEI (1996) showed that the D_A distance is generally best for inferring the correct topology for both the infinite alleles model (protein loci) and for the stepwise mutation model (microsatellites). Thus D_A was used to derive dendrograms of genetic relationships among the populations, constructed as neighbour-joining trees (SAITOU and NEI 1987). For the protein loci, most nodes were not very strongly supported, but the topology of most populations was consistent with what might be expected from their geography. The microsatellite based tree showed some similar clustering of the populations, although again very few nodes were strongly supported. Microsatellite based trees using NEI's standard (D) and REYNOLDS' distances (not shown) were very similar to each other, and had the same main clusters as the D_A tree. In contrast, the $(\delta\mu)^2$ tree showed a markedly different topology. However, nearly 80% of the pair-wise $(\delta\mu)^2$ distances were not significant, or significant only at $p < 0.05$, whereas more than 96% of the D and REYNOLDS' distances were significant at $p < 0.01$ or better. As we have no reason to suspect that either the protein or microsatellite tree is distorted (e.g. by bottlenecks of some populations – see BARKER et al. 1997b), D_A distances were computed for the combined set of 38 loci, and used to construct a neighbour joining tree (Fig. 2). This tree shows better consistency with geography and higher bootstrap values than either of the microsatellite or protein-based trees.

The WEITZMAN diversity tree based on the microsatellite D_A distances (Fig. 3), and the marginal losses of diversity for each population (Table 6) show that the losses of diversity due to extinction of any population would be highest for Ujung Pandang and Australia, and lowest for Sri Lanka South and Sarawak.

Discussion

As expected, the microsatellite loci showed many more alleles/locus and higher genetic diversity than the protein loci. Average numbers of alleles/locus were 11.8 (range 5–18) for microsatellites and 3.0 (range 2–4) for protein loci. Average gene diversities were 0.520 (range 0.259–0.702) for microsatellites and 0.339 (range 0.117–0.572) for protein loci. A high proportion of the locus-population combinations showed significant departures from

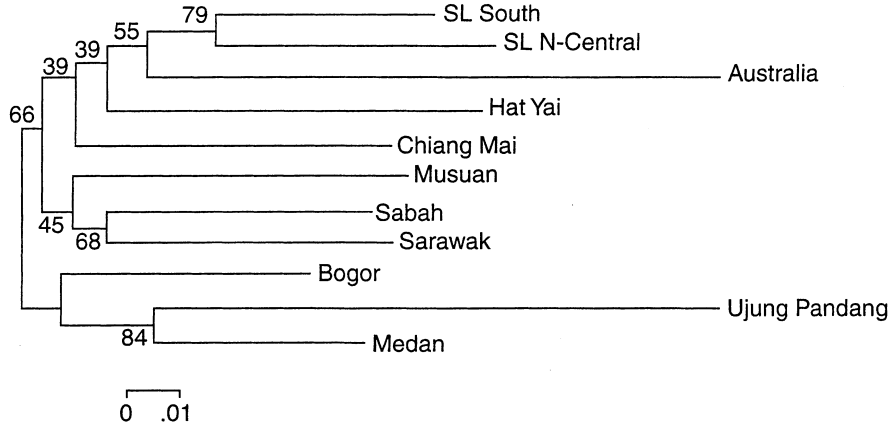

Fig. 2. Dendrogram of relationships among 11 goat populations, using D_A genetic distances and the neighbour-joining method of clustering – based on 38 loci (16 protein and 22 microsatellite). Numbers on the nodes are percentage bootstrap values from 1000 replications of resampled loci

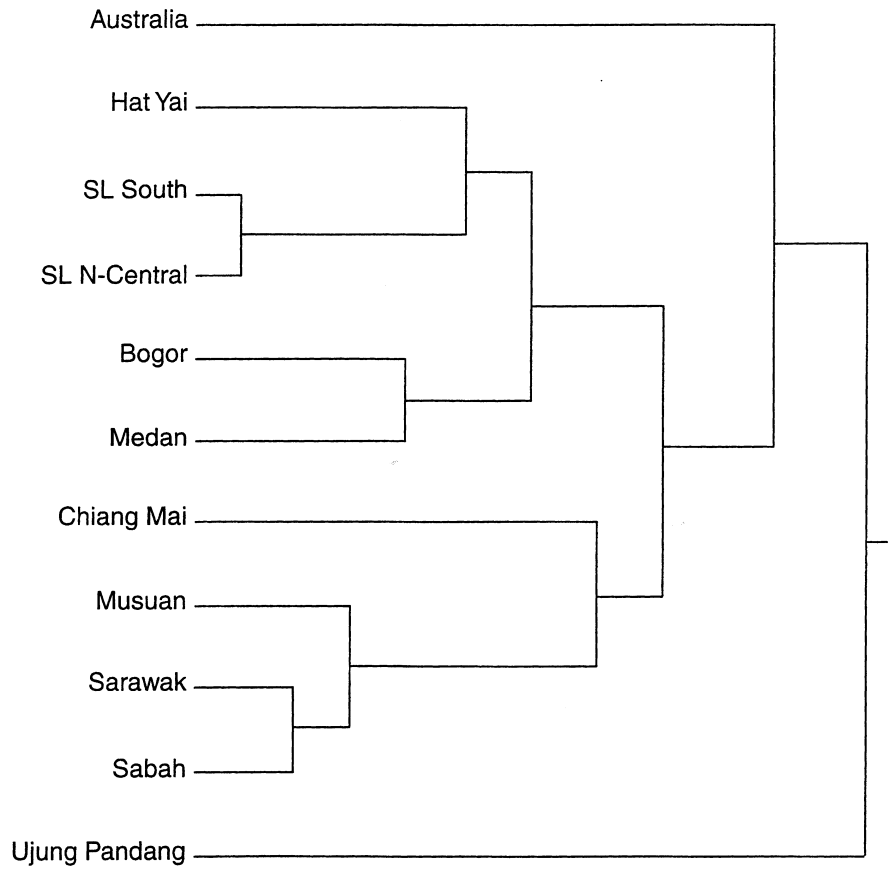

Fig. 3. Diversity dendrogram for 11 goat populations, derived using WEITZMAN's (1992) method and the pair-wise D_A genetic distances among the populations

Table 6. **Marginal losses of Weitzman diversity that would be due to extinction of each population**

Population	Marginal loss (%)
Ujung Pandang	17.66
Australia	13.56
Chiang Mai	9.43
Hat Yai	9.18
Medan	8.82
Sri Lanka N-Central	8.42
Sabah	8.33
Bogor	8.21
Musuan	8.14
Sarawak	7.69
Sri Lanka South	6.22

Hardy–Weinberg equilibrium (34.7% for protein and 21.7% for microsatellite loci), with only one significant case of an observed excess of heterozygotes for each class of marker (Table 2). Consequently, average observed heterozygosity was less than expected for all

populations, for both protein and microsatellite loci (Table 1), and mean f (F_{IS}) estimates were extremely high, ranging from 0.095 (Hat Yai) to 0.407 (Australia) for protein loci, and 0.090 (Medan) to 0.333 (Ujung Pandang) for microsatellites (Table 4), with overall means for both protein loci (0.256) and for microsatellite loci (0.230) significantly different from zero (Table 5). These high f estimates contrast strongly with most studies of livestock populations, where f generally is not significantly different from zero, even for rare breeds (e.g. ZANOTTI CASATI et al. 1988; CIAMPOLINI et al. 1995; BEHARA et al. 1998; CAÑON et al. 2000). However, significant heterozygote deficiencies have been reported in some studies. LUIKART et al. (1999) found a significant deficit of heterozygotes at one microsatellite locus (of 22) in all four goat breeds studied, and at another locus for three of the breeds, with a global deficit of heterozygotes across loci. For any single locus, however, the deficit was generally slight and not significant. For six of 15 cattle breeds, LOFTUS et al. (1999) found significant deviations ($p < 0.01$) from Hardy–Weinberg equilibrium across all 20 microsatellite loci, due to heterozygote deficiency. In both of these studies, the heterozygote deficiencies were ascribed to Wahlund effects, due to pooling samples (within breeds) from different breeding units (i.e. subpopulation structure).

The extreme heterozygote deficiencies found in these south-east Asian native goat populations could be due to any one or more of the following: segregation of nonamplifying (null) alleles, Wahlund effects, scoring bias (heterozygotes scored incorrectly as homozygotes), selection against heterozygotes or inbreeding. Distinguishing among these generally is difficult (CHRISTIANSEN et al. 1974). However, protein loci with null alleles known to be segregating are not included in the above estimates, while null alleles at other loci could not have been at high enough frequencies to account for the deficits. In any case, nulls are most unlikely to be segregating at most other protein and all microsatellite loci. Similarly, possible Wahlund effects (localities with known subpopulations) do not account for more than 40% of the observed heterozygote deficits, and then only for some loci in a few populations. Scoring bias or selection against heterozygotes may be possible for a few loci, but not for essentially all loci. Overall then, it seems that there is strong inbreeding within these south-east Asian native goat populations, presumably resulting in some way from the uncontrolled and indiscriminate mating (DEVENDRA and NOZAWA 1976), leading to small effective population sizes, breeding between relatives and consequent genetic drift. This is in direct contrast to the results of LUIKART et al. (1999) and LOFTUS et al. (1999) noted previously.

Inbreeding levels in the Australian population, which had the highest average f for protein loci (0.407) and second highest for microsatellites (0.227) are more difficult to reconcile. The animals sampled were derived from a very large feral population which has a very mixed ancestry – indigenous goats from South Africa and present day Bangladesh, imported to the early European settlement in Australia (late 1700s), Angora and Cashmere imported in the early to mid-1800s (PARSONSON 1998 and, personal communication), and most likely a number of breeds from England. We have no knowledge of the proportion of genes in the feral population contributed by each of these ancestral breeds, but it is this diverse ancestry which is no doubt reflected in the Australian population having the highest expected heterozygosity for microsatellite loci (Table 1). However, the common factor in the Australian feral and the Asian populations may well be 'uncontrolled mating'. In feral goat populations, and presumably in the uncontrolled Asian populations, breeding groups most likely comprise a dominant male and some number of females. Many males may aggregate in the vicinity of oestrus females, but the dominant male generally excludes subordinate males, and presumably sires most of the offspring (O'BRIEN 1988). With philopatric females, breeding groups will become genetically differentiated (CHESSER 1991), leading to Wahlund effects at our level of sampling (i.e. sampling at random from the whole population). In addition, the breeding groups will be expected to be inbred, with the unequal sex ratio of breeding animals and variance in sire family size causing the effective population size to be small and inbreeding to accumulate. If this mating structure is the

primary factor, then inbreeding should be less in closely managed breeds. Few data are available, but for four Spanish goat breeds and data on three protein loci (BARBANCHO et al. 1984), we compute an overall *f*-value that is not significantly different from zero (0.007 ± 0.011).

GAFFNEY et al. (1990) argued that if heterozygote deficiencies were the result of non-random mating, they should be consistent in magnitude, i.e. not show significant heterogeneity among loci within a population. HOULE (1994) advised caution in making this argument, because of sampling variation in *f*-values, variation in true *f*-values owing to finite population size and possible selection obscuring evidence of inbreeding. For the protein loci here, all populations showed highly significant heterogeneity among loci (Table 3). Although heterogeneity among populations for each locus was significant for eight of the nine loci, the magnitude of the χ^2 values was generally less (Table 3). That is, deviations from Hardy–Weinberg equilibrium for individual loci tend to be more consistent across populations than across loci within each population. In contrast, heterogeneity among loci and among populations were of similar magnitude for the microsatellite loci. Thus the more consistent deviations across populations for the protein loci suggest that selection has affected allele and genotype frequencies at some of them.

Detecting selection

The difference between the protein and microsatellite loci in the heterogeneity of deviations from Hardy–Weinberg, and in the rankings of populations for observed heterozygosity indicates something other than drift or inbreeding has affected one or other class of loci. Comparison of geographic variation in allele frequencies for different classes of polymorphic loci is a powerful method for detecting effects of natural selection (MCDONALD 1994). As the effects of drift and migration are the same for all alleles at all neutral polymorphic loci, all such alleles have the same expected F_{ST}, and there should be no significant heterogeneity of F_{ST} estimates. BARKER et al. (1997b) used the modified LEWONTIN and KRAKAUER (1973) test of POGSON et al. (1995) to compare the mean F_{ST} and the variance of F_{ST} estimates from the two classes of loci, treating all polymorphisms as two-allele polymorphisms, by using the frequencies of the overall most common allele and pooling the others. Thus unweighted F_{ST} (as computed by BIOSYS-1) were estimated from the collapsed two-allele data for each of the microsatellite and protein loci. Mean F_{ST} and variance were 0.1904 and 0.002657 for microsatellites, and 0.1227 and 0.011763 for protein loci. Applying the modified LEWONTIN-KRAKAUER tests, the means are not significantly different, but the variances are ($\chi^2_{(10)} = 44.272$, $p < 0.001$). Thus the 16 protein loci are more heterogeneous than expected from the variation shown by the microsatellite loci, indicating (assuming the microsatellite loci are neutral) selection affecting some of the protein loci. A rough test for this selection is to treat the ratio of the F_{ST} for each protein locus to the mean F_{ST} for microsatellites as a variance ratio *F*-test, and to do a two-tailed test of significance (d.f. = 10, 10). On this test, six protein loci show a significantly lower F_{ST} estimate (Alb, Amy, Mdh, Np, Pt-Tf and Xp), suggesting some form of selection affecting genetic variation at these loci. Results of the direct tests confirmed selection affecting allele frequencies at Amy, Np and Xp in most populations, and in two to five populations for Ca, Dia_1–2, Lap-1 and Mdh, but not for Alb and Pt-Tf. In all cases of significant deviation from neutrality expectations, there was an excess of genetic diversity for the number of alleles observed, consistent with selection favouring heterozygotes.

Population differentiation and history

Significant differentiation among populations is shown by θ estimates (Table 5), with similar mean estimates for protein and microsatellite loci. This differentiation undoubtedly reflects the geography of South-East Asia, with its vast archipelago of islands through

Indonesia and the Philippines, and the spread of goats throughout the region. DEVENDRA and NOZAWA (1976) postulate this spread from the centre of domestication in western Asia, through India, and possibly reaching South-East Asia by 2000 years ago. However, they quote YAMANE (1943) as suggesting that goats entered Indonesia only 'after the Moslem invasion (fourteenth century).' Our data support this latter suggestion. Using NEI's standard genetic distances from the microsatellite data, and assuming an average microsatellite mutation rate of 1.2×10^{-3} (WEBER and WONG 1993), the average divergence time between the two Sri Lankan populations (as representing the Indian subcontinent) and Medan and Bogor in Indonesia, is 64 generations (say 250 years). For a mutation rate of 5.6×10^{-4} (WEBER and WONG 1993 – dinucleotide repeats), the time is 138 generations (550 years). Given the traditional communication routes from western Asia and the Indian subcontinent to South-East Asia, introductions would have continued over time, so the estimated divergence times above would be conservative of the time of first introduction. Goats may have entered Indonesia directly from western Asia or through the Indian subcontinent, but the Musuan/Sabah/Sarawak cluster (Fig. 2) indicates a possible separate dispersion via China, as suggested previously for the spread of swamp buffalo through South-East Asia (BARKER et al. 1997b; LAU et al. 1998).

After the initial spread through the region, continued movement of animals between populations would cause the true relationships to be reticulate rather than strictly bifurcating, most likely contributing to the low bootstrap support for some of the nodes (Fig. 2). We note, however, that for pair-wise population comparisons of genotypic differentiation across all loci, all were highly significant ($p < 0.001$) for microsatellites, and all but four for protein loci. Further, the dendrogram of relationships among the south-east Asian goat populations, based on the combined set of 38 loci (Fig. 2), is consistent with their geography and the most likely pathways of dispersal through the region. The close relationship of the Australian feral population to the Sri Lankan populations indicates that the early goat introductions to Australia from the Indian subcontinent and nearby regions have made a major contribution to the genetic composition of the Australian population.

Although the microsatellite-based phylogeny and the Weitzman diversity tree (Fig. 3) showed some identical clusters, it should not be expected that they will give the same information (THAON D'ARNOLDI et al. 1998). But as branch lengths in the neighbour joining phylogeny indicate the magnitude of the differences between populations, one might expect a positive relationship between branch length for a population and expected loss of diversity on extinction of that population. Both the microsatellite phylogeny and the Weitzman diversity tree were derived from the D_A distances, and for these goat populations, the regression coefficient for diversity loss on branch lengths was 0.779 ± 0.076 ($p < 0.001$). Thus the branch lengths in the neighbour joining D_A phylogeny apparently give essentially the same information as the Weitzman analysis on expected loss of diversity due to extinction of individual populations. This may not be true for other sets of populations, and of course, the Weitzman analysis allows estimation of expected loss of diversity if two or more breeds were to become extinct.

In addition to the phylogenetic relationships among the populations, we note the long branch lengths for Australia and Ujung Pandang in the microsatellite tree (not shown) and in the combined tree (Fig. 2). These two populations also were identified by the Weitzman analysis (Fig. 3, Table 6) as those whose extinction would cause the greatest loss of diversity. They are exceptional in a number of ways – higher expected heterozygosity and higher proportion of private alleles than other populations (9.9% of alleles in Ujung Pandang are unique to that population, and 5.2% for Australia), but also highest inbreeding coefficients (F_{IS} – Table 6). Although there were relatively few cases of significant linkage disequilibrium overall, these two populations contrast in that Ujung Pandang showed more than any other population, while Australia had none. Yet the Australian feral population, because of its recent and diverse breed ancestry, might have been expected to exhibit higher linkage disequilibrium than other populations. The long branch lengths for these two

populations and the significant linkage disequilibrium for Ujung Pandang are compatible with the high inbreeding coefficients (F_{IS}). With such high inbreeding, however, loss of low frequency alleles and reduced observed heterozygosity would be expected, yet these populations have above average to very high expected heterozygosity for both protein and microsatellite loci.

Acknowledgements

We gratefully acknowledge the cooperation and assistance during sample collection of Professor C. Chantalakhana, Kasetsart University, Dr B. Cheva-Isarakul, Chiang Mai University and Ms. Ancharlee, Department of Livestock Development, Thailand; Professor R. L. Anacleto, Central Mindanao University of the Philippines; Dr P. Abeynayake, University of Peradeniya, Sri Lanka; Dr Sudarisman and Dr P. Daniels, Balitvet, Indonesia; Dr S. P. Chia, Department of Animal Services, Sabah; Dr K. Y. Hsiung, Department of Agriculture, Sarawak; Dr M. Hilmi, Universiti Putra Malaysia; and Dr B. Restall, NSW Department of Agriculture and Fisheries, Wollongbar, Australia. We are grateful to Ms Rohani Man and Ms Norliza Osman for laboratory assistance, and to Dr J.-L. Foulley for carrying out the Weitzman analysis. This work was supported by the Australian Centre for International Agricultural Research (Project no. 8364 and Contract AS1 : 96/66), the Malaysian Program for Intensification of Research in Priority Areas (Project no. 1–07–05–057), Universiti Putra Malaysia and the University of Malaya.

References

Barbancho, M.; Llanes, D.; Morera, L.; Garzón, R.; Rodero, A., 1984: Genetic markers in the blood of Spanish goat breeds. Anim. Bld Groups Biochem. Genet. **15:** 207–212.

Barker, J. S. F.; Tan, S. G.; Selvaraj, O. S.; Mukherjee, T. K., 1997a: Genetic variation within and relationships among populations of water buffalo (*Bubalus bubalis*). Anim. Genet. **28:** 1–13.

Barker, J. S. F.; Moore, S. S.; Hetzel, D. J. S.; Evans, D.; Tan, S. G.; Byrne, K., 1997b: Genetic diversity of Asian water buffalo (*Bubalus bubalis*): microsatellite variation and a comparison with protein-coding loci. Anim. Genet. **28:** 103–115.

Behara, A. M. P.; Colling, D. T.; Cothran, E. G.; Gibson, J. P., 1998: Genetic relationships between horse breeds based on microsatellite data: applications for livestock conservation. Proceedings 6th Wld. Cong. Genet. Appl. Livestock Prod. **28:** 119–122.

Cañon, J.; Checa, M. L.; Carleos, C.; Vega-Pla, J. L.; Vallego, M.; Dunner, S., 2000: The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. Anim. Genet. **31:** 39–48.

Chesser, R. K., 1991: Gene diversity and female philopatry. Genetics **127:** 437–447.

Christiansen, F. B.; Frydenberg, O.; Gyldenholm, A. O.; Simonsen, V., 1974: Genetics of Zoarces populations VI. Further evidence, based on age group samples, of a heterozygote deficit in the ESTIII polymorphism. Hereditas **77:** 225–236.

Ciampolini, R.; Moazami-Goudarzi, K.; Vaiman, D.; Dillmann, C.; Mazzanti, E.; Foulley, J.-L.; Leveziel, H.; Cianci, D., 1995: Individual multilocus genotypes using microsatellite polymorphisms to permit the analysis of the genetic variability within and between Italian beef cattle breeds. J. Anim. Sci. **73:** 3259–3268.

Devendra, C.; Nozawa, K., 1976: Goats in South East Asia – their status and production. Z. Tierz. Züchtbiol. **93:** 101–120.

Estoup, A.; Angers, B., 1998: Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. In: Carvalho, G. ed., Advances in Molecular Ecology. IOS Press, Amsterdam, pp. 55–86.

FAO., 1995: World Watch List for Domestic Animal Diversity, 2nd edn., FAO, Rome, Italy.

Food and Fertilizer Technology Center, 1984: Goat production in southeast Asia. Extension Bull. **206:** 44pp.

Gaffney, P. M.; Scott, T. M.; Koehn, R. K.; Diehl, W. J., 1990: Interrelationships of heterozygosity, growth rate and heterozygote deficiencies in the coot clam, *Mulinia lateralis*. Genetics **124:** 687–699.

Goldstein, D. B.; Linares, A. R.; Cavalli-Sforza; L. L.; Feldman, M. W., 1995: Genetic absolute dating based on microsatellites and the origin of modern humans. Proc. Natl. Acad. Sci. USA **92:** 6723–6727.

Goodman, S. J., 1997: R_{ST} Calc: 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., 1995: FSTAT (Version 1.2): a computer program to calculate F-statistics. J. Hered. **86:** 485–486.

HARRIS, H.; HOPKINSON, D. A., 1976: Handbook of Enzyme Electrophoresis in Human Genetics. North-Holland Publishing Co., Amsterdam.
HEDRICK, P. W., 1999: Perspective: Highly variable loci and their interpretation in evolution and conservation. Evolution **53:** 313–318.
HIROOKA, H.; MUKHERJEE, T. K.; PANANDAM, J. M.; HORST, P., 1997: Genetic parameters for growth performance of the Malaysian local goats and their crossbreds with the German (improved) Fawn goats. J. Anim. Breed. Genet. **114:** 191–199.
HOCHBERG, Y., 1988: A sharper Bonferroni procedure for multiple tests of significance. Biometrika **75:** 800–802.
HOULE, D., 1994: Adaptive distance and the genetic basis of heterosis. Evolution **48:** 1410–1417.
KATSUMATA, M.; AMANO, T.; TANAKA, K.; NOZAWA, K.; BAHK, K.; PARK, B.; LEE, C., 1982: Blood protein variations of the Korean native goats. Jpn. J. Zootech. Sci. **53:** 521–527.
LAU, C. H.; DRINKWATER, R. D.; YUSOFF, K.; TAN, S. G.; HETZEL, D. J. S.; BARKER, J. S. F., 1998: Genetic diversity of Asian water buffalo (*Bubalus bubalis*): Mitochondrial DNA D-loop and cytochrome b sequence variation. Anim. Genet. **29:** 253–264.
LESSIOS, H. A., 1992: Testing electrophoretic data for agreement with Hardy–Weinberg expectations. Marine Biol. **112:** 517–523.
LEWIS, P. O.; ZAYKIN, D., 1999: Genetic Data Analysis: computer program for the analysis of allelic data, Version 1.0 (d12). Free program distributed by the authors over the internet from the GDA Home Page at http://chee.unm.edu/gda/.
LEWONTIN, R. C.; KRAKAUER, J., 1973: Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics **74:** 175–195.
LI, C. C., 1976: First Course in Population Genetics. Boxwood Press. Pacific Grove, California.
LOFTUS, R. T.; ERTUGRUL, O.; HARBA, A. H.; EL-BARODY, M. A. A.; MACHUGH, D. E.; PARK, S. D. E.; BRADLEY, D. G., 1999: A microsatellite survey of cattle from a centre of origin: the Near East. Mol. Ecol. **8:** 2015–2022.
LUIKART, G.; BIJU-DUVAL, M.-P.; ERTUGRUL, O.; ZAGDSUREN, Y.; MAUDET, C.; TABERLET, P., 1999: Power of 22 microsatellite markers in fluorescent multiplexes for parentage testing in goats (*Capra hircus*). Anim. Genet. **30:** 431–438.
LYNCH, M.; MILLIGAN, B. G., 1994: Analysis of population genetic structure with RAPD markers. Mol. Ecol. **3:** 91–99.
MCDONALD, J. H., 1994: Detecting natural selection by comparing geographic variation in protein and DNA polymorphisms. In: GOLDING, B. ed., Non-Neutral Evolution Theories and Molecular Data. Chapman & Hall, New York, pp. 88–100.
MANTEL, N., 1967: The detection of disease clustering and a generalized regression approach. Cancer Res. **27:** 209–220.
MINCH, E.; RUIZ-LINARES, A.; GOLDSTEIN, D. B.; FELDMAN, M. W.; CAVALLI-SFORZA, L. L., 1995: Microsat (Version 1.5d): a computer program for calculating various statistics on microsatellite allele data. WWW: http://lotka.stanford.edu/microsat.html.
MOORE, S. S.; BYRNE, K.; MCCARTHY, F.; BARENDSE, W.; WOMACK, J. E.; HETZEL, D. J. S., 1994: Characterisation of 65 bovine microsatellites. Mamm. Genome **5:** 84–90.
NEI, M., 1973: Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA **70:** 3321–3323.
NEI, M., 1978: Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89:** 583–590.
NEI, M.; TAJIMA, F.; TATENO, Y., 1983: Accuracy of estimated phylogenetic trees from molecular data II. Gene frequency data. J. Mol. Evol. **19:** 153–170.
NOZAWA, K.; SHINJO, A.; SHOTAKE, T., 1978: Population genetics of farm animals III. Blood-protein variations in the meat goats in Okinawa Islands of Japan. Z. Tierz. Züchtbiol. **95:** 60–77.
O'BRIEN, P. H., 1988: Feral goat social organization: a review and comparative analysis. Appl. Anim. Behav. Sci. **21:** 209–221.
PANANDAM, J. M.; MUKHERJEE, T. K.; HORST. P., 1987: Comparison of Katjang goats and their first and second generation crosses with German Fawn – individual and maternal heterosis for growth traits. SABRAO J. **19:** 61–68.
PARSONSON, I. M., 1998: The Australian Ark. CSIRO Publishing, Collingwood.
POGSON, G. H.; MESA, K. A.; BOUTILIER, R. G., 1995: Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. Genetics **139:** 375–385.
RAYMOND, M.; ROUSSET, F., 1995: GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. J. Hered. **86:** 248–249.
REYNOLDS, J.; WEIR, B. S.; COCKERHAM, C. C., 1983: Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics **105:** 767–779.
RICHARDSON, B. J.; BAVERSTOCK, P. R.; ADAMS, M., 1986: Allozyme Electrophoresis. A Handbook for Animal Systematics and Population Studies. Academic Press Australia, Sydney.

ROUSSET, F., 1997: Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics **145:** 1219–1228.
SAITOU, N.; NEI, M., 1987: The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406–425.
SCHNEIDER, S.; ROESSLI, D.; EXCOFFIER, L., 2000: Arlequin, Version 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory. University of Geneva, Switzerland.
SLATKIN, M., 1995: A measure of population subdivision based on microsatellite allele frequencies. Genetics **139:** 457–462.
SOKAL, R. R.; ROHLF, F. J., 1981: Biometry, 2nd edn. W.H. Freeman & Co., San Francisco.
SWOFFORD, D. L.; SELANDER, R. B., 1989: BIOSYS-1 A computer program for the analysis of allelic variation in population genetics and biochemical systematics. Release 1.7. Illinois Natural History Survey, Campaign, Ill.
TAKEZAKI, N.; NEI, M., 1996: Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics **144:** 389–399.
TAN, S. G.; SELVARAJ, O. S.; SEKARAN, M.; MUKHERJEE, T. K.; BARKER, J. S. F., 1990: Biochemical polymorphisms in swamp buffaloes (*Bubalus bubalis*) I. Cellulose acetate gel methodology. In: VIDYADARAN, M. K.; AZMI, T. I.; BASRUR, P. K. (eds.) Seminar on Buffalo Genotypes for Small Farms in Asia, 1989. Universiti Pertanian Malaysia, Serdang, pp. 255–261.
TAN, S. G.; SELVARAJ, O. S.; ROSLIN, J.; AKLIMA, O.; GAN, Y. Y.; MUKHERJEE, T. K.; BARKER, J. S. F., 1991: Electrophoretic studies on southeast Asian buffalo and goats: Methodology. In: TULLOH, N. H.; ed. Buffalo and Goats in Asia: Genetic Diversity and its Application. ACIAR Proceedings 34, Australian Centre for International Agricultural Research, Canberra, pp. 28–33.
TAN, S. G.; BARKER, J. S. F.; SELVARAJ, O. S.; MUKHERJEE, T. K.; WONG, Y. F., 1993: Genetic studies of water buffalo blood markers. I. Red cell acid phosphatase, albumin, catalase, red cell α-esterase-3, group-specific component and protease inhibitor. Biochem. Genet. **31:** 223–230.
THAON D'ARNOLDI, C.; FOULLEY, J.-L.; OLLIVIER, L., 1998: An overview of the Weitzman approach to diversity. Genet. Sel. Evol. **30:** 149–161.
VAIMAN, D.; SCHIBLER, L.; BOURGEOIS, F.; OUSTRY, A.; AMIGUES, Y.; CRIBIU, E. P., 1996: A genetic linkage map of the male goat genome. Genetics **144:** 279–305.
WEBER, J. L.; WONG, C., 1993: Mutation of human short tandem repeats. Hum. Mol. Genet. **2:** 1123–1128.
WEIR, B. S., 1990: Genetic Data Analysis. Sinauer, Sunderland.
WEIR, B. S.; COCKERHAM, C. C., 1984: Estimating *F*-statistics for the analysis of population structure. Evolution **38:** 1358–1370.
WEITZMAN, M. L., 1992: On diversity. Q. J. Econ. **107:** 363–405.
WEITZMAN, M. L., 1993: What to preserve? An application of diversity theory to crane conservation. Q. J. Econ. **108:** 157–183.
WRIGHT, S., 1951: The genetical structure of populations. Ann. Eugenics (London) **15:** 323–354.
WRIGHT, S., 1978: Evolution and the Genetics of Populations, Vol. 4. Variability Within and Among Natural Populations. University of Chicago Press, Chicago.
YAMANE, J., 1943: Animal Husbandry of the East Indies. Yoken-do, Tokyo (in Japanese).
YANG, L.; ZHAO, S. H.; LI, K.; PENG, Z. Z.; MONTGOMERY, G. W., 1999: Determination of genetic relationships among five indigenous Chinese goat breeds with six microsatellite markers. Anim. Genet. **30:** 452–455.
ZANOTTI CASATI, M.; GANDINI, G. C.; LEONE, P.; ROGNONI, G., 1988: Genetic relationship among four sheep breeds of the Italian Alpine Ark. J. Anim. Breed. Genet. **105:** 135–142.

Author's address: J. S. F. BARKER, School of Rural Science and Natural Resources, University of New England, Armidale, NSW 2351, Australia. E-mail: sbarker@metz.une.edu.au

Appendix 1. **Enzymes or proteins and electrophoretic assay systems**

	Enzyme (or protein)	EC number	Source*	System**	Polymor phic***
Acp	Acid phosphatase	3.1.3.2	RBC	CAE	–
Ak	Adenylate kinase	2.7.4.3	RBC	CAE	–
Alb	Albumin	–	Plasma	CAE	+
Amy	Amylase	3.2.1.1	Plasma	CAE	+
c Amy-1	Cathodal amylase-1	3.2.1.1	Plasma	AGE	+
c Amy-2	Cathodal amylase-2	3.2.1.1	Plasma	AGE	–
Blvr	Biliverdin reductase	1.3.1.24	RBC	CAE	–
Ca	Carbonic anhydrase	4.2.1.1	RBC	CAE	+
Cat	Catalase	1.11.1.6	RBC	PAGE	+
Cp	Ceruloplasmin	–	Plasma	PAGE	–
Dia_1–1	NADH Diaphorase, zone 2	1.6.2.2	RBC	CAE	+
Dia_1–2	NADH Diaphorase, zone 3	1.6.2.2	RBC	CAE	+
Dpgm	Diphosphoglyceromutase	2.7.5.4	RBC	CAE	–
Est-D	Esterase D	3.1.1.-	RBC	CAE	–
α Est-1	α Esterase-1	3.1.1.-	Plasma	PAGE	–
α Est-3	α Esterase-3	3.1.1.-	RBC	PAGE	–
β Est-1	β Esterase-1	3.1.1.-	RBC	PAGE	–
Fdp	Fructose-1, 6-diphosphatase	3.1.3.11	RBC	CAE	–
Fk	Fructokinase	2.7.1.4	RBC	CAE	–
Fum	Fumarase	4.2.1.2	RBC	CAE	–
Gldh	Glucose dehydrogenase	1.1.1.118	WBC	CAE	–
Glo I	Glyoxalase I	4.4.1.5	RBC	PAGE	–
Got	Glutamate oxaloacetate transaminase	2.6.1.1	RBC	CAE	–
G6pdh	Glucose 6-phosphate dehydrogenase	1.1.1.49	RBC	CAE	–
Gapdh	Glyceraldehyde-3-phosphate	1.2.1.12	RBC	CAE	–
Gc	Group specific component	–	Plasma	PAGIF	+
Gpi	Glucose phosphate isomerase	5.3.1.9	RBC	CAE	–
Gpt	Glutamate pyruvate transaminase	2.6.1.2	RBC	CAE	–
Gpx	Glutathione peroxidase	1.11.1.9	RBC	PAGE	–
Gsr	Glutathione reductase	1.6.4.2	RBC	CAE	–
Hb	Haemoglobin	–	RBC	CAE	+
Hk	Hexokinase	2.7.1.1	RBC	CAE	–
Idh	Isocitrate dehydrogenase	1.1.1.42	RBC	CAE	–
Itp	Inosine triphosphatase	3.6.1.19	RBC	PAGE	–
Lap-1	Leucine aminopeptidase-1	3.4.11.1	Plasma	PAGE	+
Lap-2	Leucine aminopeptidase-2	3.4.11.1	Plasma	PAGE	–
Ldh	Lactate dehydrogenase	1.1.1.27	RBC	CAE	–
Mdh	Malate dehydrogenase	1.1.1.37	RBC	CAE	+
Me	Malic enzyme	1.1.1.40	RBC	CAE	+
Mpi	Mannose phosphate isomerase	5.3.1.8	RBC	CAE	–
NADPH Dia-2	NADPH Diaphorase, zone 2	1.6.2.2	RBC	CAE	–
Np	Purine nucleoside phosphorylase	2.4.2.1	RBC	CAE	+
Pep A	Peptidase A	3.4.	RBC	CAE	–
Pep B	Peptidase B	3.4.	RBC	CAE	–
Pep C	Peptidase C	3.4.	RBC	CAE	–
Pep D	Peptidase D	3.4.	RBC	PAGE	–
Pep D	Peptidase D	3.4.	Plasma	PAGE	–
6 Pgdh	6-Phosphogluconate dehydrogenase	1.1.1.44	RBC	CAE	–
Pgm-2	Phosphoglucomutase-2	5.4.2.2	RBC	CAE	–
Pgp	Phosphoglycolate phosphatase	3.1.3.18	RBC	PAGE	–
Pk	Pyruvate kinase	2.7.1.40	RBC	CAE	–
Pp	Inorganic pyrophosphatase	3.6.1.1	RBC	PAGE	–
Sod	Superoxide dismutase	1.15.1.1	RBC	CAE	–
Sordh	Sorbitol dehydrogenase	1.1.1.14	RBC	CAE	–
Tf	Transferrin	–	Plasma	CAE	+

Appendix 1. **continued**

	Enzyme (or protein)	EC number	Source*	System**	Polymor phic***
Pt-Tf	Post-transferrin	–	Plasma	PAGE	+
Tpi	Triose phosphate isomerase	5.3.1.1	RBC	CAE	–
Umpk	Uridine monophosphate kinase	2.7.1.48	RBC	CAE	–
Xp	X Protein	–	RBC	STAGE	+

*RBC = Red blood cells, WBC = White blood cells. **CAE = cellulose acetate electrophoresis, PAGE = polyacrylamide electrophoresis, STAGE = starch gel electrophoresis, PAGIF = polyacrylamide gel, isoelectric focussing. ***+ indicates those loci that are polymorphic in one or more population.

Appendix 2. **Microsatellite loci investigated, running conditions, allele size ranges, number of alleles and expected heterozygosity**

Locus	Co-load group	Annealing Temp.	Dye	Size range	No. of alleles	Expected heterozygosity	
						Range	Mean
BL25	1	60	6FAM	173–190	6	0.000–0.554	0.278
INRA003	1	60	TET	182–193	9	0.043–0.569	0.266
TGLA325	1	55	TET	104–130	13	0.392–0.866	0.622
TGLA245	1	55	6FAM	121–147	8	0.048–0.615	0.292
RM004	1	55	HEX	107–133	10	0.206–0.653	0.463
RME23	2	55	6FAM	139–155	9	0.152–0.763	0.517
INRA071	2	50	TET	182–200	12	0.000–0.801	0.356
TGLA378	2	55	TET	106–128	11	0.529–0.824	0.682
INRA005	2	50	HEX	134–148	8	0.000–0.697	0.461
TGLA179	2	55	HEX	67–95	18	0.410–0.866	0.710
INRA177	3	55	HEX	167–201	18	0.246–0.859	0.693
CSSM026	3	55	6FAM	256–281	12	0.129–0.852	0.564
TGLA272	3	55	6FAM	92–128	15	0.333–0.692	0.515
OarFCB48	3	60	TET	145–174	14	0.193–0.815	0.607
MAF070	4	60	HEX	137–163	13	0.554–0.886	0.721
RM188	4	60	6FAM	115–146	13	0.482–0.824	0.664
INRA063	4	50	TET	156–181	13	0.000–0.833	0.475
BMC1009	4	60	TET	280–299	11	0.351–0.790	0.663
BM3205	5	50	TET	211–233	12	0.431–0.889	0.715
BM757	5	50	HEX	176–192	7	0.000–0.411	0.064
RM006	5	55	6FAM	122–147	13	0.434–0.901	0.706
ILSTS005	5	55	TET	173–186	5	0.048–0.605	0.299
RM044	5	60	HEX	77–115	15	0.345–0.841	0.662
TGLA75	6	55	TET	156–200	13	0.100–0.807	0.508
OarFCB193	6	55	6FAM	114–136	17	0.167–0.813	0.574