



Genetic diversity and population structure of Zackel sheep and other Hungarian sheep breeds

V. Neubauer¹, C. Vogl¹, J. Seregi², L. Sáfár³, and G. Brem¹

¹Institute for Animal Breeding and Genetics, Department for Biomedical Science, University of Veterinary Medicine Vienna, Vienna, Austria

²Kaposvár University, Kaposvár, Hungary

³Hungarian Sheep and Goat Breeders' Association, Budapest, Hungary

Correspondence to: C. Vogl (claus.vogl@vetmeduni.ac.at)

Received: 24 November 2014 – Revised: 11 June 2015 – Accepted: 21 July 2015 – Published: 31 August 2015

Abstract. More than 6000 samples of 18 different Hungarian sheep breeds were genotyped for 10 unlinked microsatellite loci. After data cleaning, 5434 sheep remained in the analysis. Some locus–breed combinations show deviation from Hardy–Weinberg equilibrium, possibly due to null alleles or the Wahlund effect. All breeds show high genetic variability; the lowest expected heterozygosity is that of the British Milkshope (He = 0.588). The Transylvanian Zackel has the highest expected heterozygosity (He = 0.790). It is also the least differentiated breed (Fst = 0.020). Pairwise genetic distances among breeds range from 0.021 between Hungarian Merino and both Mutton Merino and Transylvanian Zackel to greater than 0.105 between British Milkshope and all other breeds. The three Zackel populations – white, black, and Transylvanian Zackel – show small genetic distances among each other, with pairwise Fst values from 0.030 to 0.058. The Transylvanian Zackel tended to have close relationships to some other breeds too, probably due to its low differentiation. Given the individual genotypic information, a Bayesian analysis assigned individuals to breeds generally correctly.

1 Introduction

The Hungarian Zackel sheep originates from the Hungarian lowlands. During the Middle Ages and thereafter until about 1800, the Zackel was the most popular sheep breed in the Pannonian area. Hybrids never established against pure-blooded herds. During the 19th century, however, the Merino sheep slowly replaced the Zackel in the lowlands. In the second half of the 20th century, pure Zackel mainly survived in the national gene reserve in the Hortobágy Pusztá (Berger and Fischerleitner, 2009).

In 1983, the Hungarian Zackel sheep breeder association was founded to maintain old breeds and to preserve genetic diversity. The restoration of the breed was successful in Hungary; currently, pure Zackel number about 6700 individuals (DAD-IS, 2015). The Hungarian Zackel are fertile, vital, robust, and adaptable, and are therefore nowadays used for landscape conservation of extensively cultivated pastures in the Pannonian area (Berger and Fischerleitner, 2009). To-

day two colour types, black and white, are bred strictly separated. Characteristic for both colour types are the screw-shaped horns, which are held in a low V.

Another Zackel breed is the Transylvanian Zackel sheep. It originated in Transylvania and has been known since the end of the Middle Ages. Today, purebred Transylvanian Zacks are only found in the Carpathian Mountains. In Romania, the breed is still kept in large numbers and therefore not protected. In Hungary, a small population is maintained as a genetic resource. Unlike the Hungarian Zackel, male Transylvanian Zackel sheep have regular, open-spiral-shaped horns. Females are usually hornless or bear short vestigial horns (Hortobágyi Nonprofit Kft., 2014).

The Babolna Tetra is another Hungarian breed that originated relatively recently via a cross of the Merino sheep with the fertile and hardy breeds Romanov and Finnish Landrace. Subsequently, the negative traits introduced by the crosses (high body weight, litter sizes of four to five lambs) were selected against while maintaining the high fertility and asea-

Table 1. Details on the 18 sheep breeds: the consecutive number (used for the illustrations), the English and Hungarian names, abbreviations, and numbers of individuals (after data cleaning).

Consecutive Number	English/German name	Hungarian name	Abbreviation	Number of individuals
1	Awassi	Awassi	<i>awa</i>	38
2	Babolna Tetra	Bábolna Tetra	<i>bab</i>	129
3	British Milkssheep	Brit tejelő	<i>brt</i>	30
4	Charollais	Charollais	<i>cha</i>	150
5	Tsigai	Cigája	<i>cig</i>	543
6	Cikta	Cikta	<i>cik</i>	146
7	Hungarian Zackel white	Hortobágyi racka fehér	<i>fehr</i>	403
8	Hungarian Zackel black	Hortobágyi racka fekete	<i>fekr</i>	330
9	Transylvanian Zackel	Gyimesi racka	<i>gyr</i>	275
10	Île-de-France	Ile de france	<i>ile</i>	454
11	Lacaune	Lacaune	<i>lac</i>	245
12	Merinolandschaf	Landschaf merinó	<i>lsm</i>	168
13	Hungarian Merino	Magyar merinó	<i>mm</i>	907
14	German Blackheaded Mutton	Német feketefejú húsjuh	<i>nff</i>	261
15	German Mutton Merino	Német húsmerinó	<i>nhm</i>	670
16	Suffolk	Suffolk	<i>suf</i>	528
17	Dairy Tsigai	Tejelő cigája	<i>tcig</i>	89
18	Texel	Texel	<i>tex</i>	68

sonality of breeding. The stock has been bred for 30 years without further intake of outside blood. At present, 800 ewes are registered.

The Tsigai breed was introduced from the Balkans to Hungary and the neighbouring countries at the end of the 18th century. It was used for milk, wool, and meat. The advent of the Merino sheep almost drove this race to extinction. It was preserved with imported sheep from neighbouring countries. Presently it numbers over 3000 ewes.

A description of the other 13 breeds investigated (see Table 1) is given in Sambraus (2011, p. 191) and websites of breeding associations (Oklahoma State University, 1995; DAD-IS, 2015).

As can be seen from the description of the breeds' history, distinctive breeds either became established from obscure sources or were established by cross-breeding and hybridisation. Thereafter the breeds were maintained distinct, generally without contributions from other breeds. At various times, genetic diversity may have suffered from bottlenecks.

For current preservation programmes, knowledge of the genetic relationships of the Zackel breeds has become important. In this study almost 6000 samples of 18 different sheep breeds from Hungarian livestock breeders were used, including three Zackel breeds: white Hungarian Zackel (*feher racka*), black Hungarian Zackel (*fekete racka*), and Transylvanian Zackel (*gyimesi racka*). Via highly polymorphic microsatellite loci, the genetic diversity among the Zackel breeds, the relationship among sheep breeds, and the assignment of individuals to sheep breeds, including the possibility of hybridisation, was inferred.

For inferring relationships among sheep breeds we used traditional pairwise F_{st} distances and a dendrogram based on UPGMA (unweighted pair-group method with arithmetic mean). The sheep breeds' population structure, however, fits the assumptions of a Bayesian clustering method (Pritchard et al., 2000) based on a probabilistic population genetic model. The algorithm is implemented in the program *structure* (Pritchard et al., 2000). It utilises prior information of an individual's origin as a bias of assignment to a breed while allowing for the possibility of hybridisation among breeds. With this method, it is thus possible to adequately represent the genetic relationships among Hungarian sheep breeds.

2 Material and methods

2.1 Samples

Samples of more than 6000 sheep were collected in Hungary as part of routine diagnostic procedures (parentage analysis). After data cleaning (see below), a total amount of 5434 unrelated individuals from 18 different breeds were investigated. The main focus of this study was on Zackel sheep, consisting of three breeds: the white and black Hungarian and the Transylvanian Zackel. Information on all breeds, including their English and Hungarian names, abbreviations used in this paper, and actual sample size, can be found in Table 1.

To confirm our methods, we used data from another group of 23 Zackel sheep.

Table 2. List of microsatellite markers used, allelic size, number of identified alleles, mean number of alleles per locus, and observed (Ho) and expected (He) heterozygosity for each locus for the whole population.

Number	Locus	Allelic size range	Mean allelic size	Number of identified alleles	Mean number of alleles per locus	Ho	He
1	HSC	14–301	162.7	23	13.8	0.823	0.827
2	INRA063	30–216	100.3	25	13.8	0.737	0.746
3	MAF214	2–269	90.7	31	10.2	0.626	0.621
4	OarAE129	4–163	133.7	14	6.4	0.572	0.653
5	OarCP49	24–134	75.2	34	16.3	0.821	0.834
6	OarFCB11	16–151	82.4	17	9.9	0.792	0.792
7	INRA005	18–156	51.4	21	12.1	0.822	0.823
8	INTRA23	18–221	154.2	20	11.4	0.813	0.809
9	MA065	10–139	70.0	14	8.1	0.676	0.692
10	McM527	8–183	78.1	15	7.9	0.761	0.757
Mean			99.9	21.4	11	0.744	0.755

2.2 Microsatellite markers

The 10 highly polymorphic, unlinked microsatellite markers used in this study are listed in Table 2. Some of the markers used are also found in the list of ISAG-FAO recommended microsatellite markers for molecular genetic characterisation of animal genetic resources (FAO, 2011). The same panel of microsatellite markers was used by, for example, Baumung et al. (2006) for analysis of Austrian sheep breeds. Several studies have shown that, with between 5 (Buchanan et al., 1994) and 12 microsatellites (MacHugh, 1994), significant differences among breeds that are not too closely related can be detected.

The microsatellites were amplified by polymerase chain reaction and scored according to their lengths as part of routine diagnostics. We removed individuals with data missing from more than six loci from the analysis. For the remaining data, we imputed alleles of monomorphic loci to either a segregating null allele or the homozygous allele according to their most probable state, assuming Hardy–Weinberg equilibrium (HWE) within a breed. For further analysis, we treated these imputed alleles as all other alleles. We chose this approach to allow the comparison of analyses between *structure* (Pritchard et al., 2000), which allows for imputation of null alleles, and *Arlequin* (Excoffier et al., 2005), which does not.

2.3 Statistical analysis

The original microsatellite allelic data are available from the authors upon request.

We used the software package *Arlequin* (Excoffier et al., 2005) for general population genetics data analysis and genetic differentiation. We determined the number of alleles and the allelic size range over all samples for each locus, without correction for null alleles. We calculated observed (Ho) and expected (He) heterozygosities for each breed and

locus combination and then averaged the results over loci to obtain an estimate per breed. We performed exact locus-by-locus tests for HWE, with simulations of 10^6 steps in Markov chain repeats and 10^5 dememorisation steps for each locus in each of the 18 populations (Guo and Thompson, 1992); the p values were estimated from the simulations.

As a measure of the genetic distance between the breeds, we determined F_{st} for all pairs of populations, as well as Reynolds' genetic distance (RD), a transformed pairwise F_{st} (Reynolds et al., 1983). Distances based on the co-ancestry coefficient, like RD, are designed to measure the divergence between populations caused by drift. Therefore pairwise F_{st} and RD are appropriate for short-term evolution when mutation can be neglected (Excoffier et al., 2005). The matrix of pairwise F_{st} distances was used as a dissimilarity matrix to produce a dendrogram using UPGMA.

We used the program *structure* (Pritchard et al., 2000) to infer population structure. The method is based on a Bayesian model to infer the population assignment of individuals including the possibility of hybrids. We used the USEPOPINFO model, which assumes that the predefined populations are usually correct. We conducted several runs with slightly different model parameters.

For the final runs, parameters recommended in Pritchard et al. (2010) were used. Due to our extremely large data set we decided on a burn-in period of 2×10^4 and 2×10^4 repeats in order to avoid run times that are too long. We checked consistency by performing several runs with the same parameters. The parameter “Migprior”, the probability that a specific individual is an immigrant to the population, or has an immigrant ancestor in the last G generations, was set to $v = 0.05$. For this parameter, values in the range of 0.001–0.1 are suggested (Pritchard et al., 2000).

Some breeds show significant deviations from HWE (see results), which may be caused by population subdivision within these breeds, i.e. the Wahlund effect (Hartl and Clark,

2007). Hence, we estimated with the ADMIXTURE model in *structure* the most likely number of partitions for single breeds, independent of breed affiliation. The ADMIXTURE model assumes that each individual draws some fraction of its genome from each of the K populations. The output records the *posterior mean estimates* of these proportions. Conditional on the ancestry vector, the origin of each allele is independent (Pritchard et al., 2000). To estimate the most likely number of K subpopulations, we conducted several runs with models from $K = 1$ to $K = 4$ subpopulations, with repeats of 10^5 for both burn-in and Markov chain Monte Carlo. Then we compared the log-likelihood of each model with the harmonic mean estimator (Pritchard et al., 2000) to evaluate the most likely value of K .

Moreover, the genotypes of the 23 additional Hungarian Zackel sheep were used for assignment to sheep breed. For this task, the predefined population was used for all other individuals, while the 23 individuals were assigned to the breeds. We conducted runs with 2×10^4 iterations for burn-in and the sampler for this task (Pritchard et al., 2000).

3 Results

3.1 Microsatellite loci

A total of 215 alleles were found across the 10 loci. The FAO (1998) recommends using only loci with a minimum number of four alleles for genetic distance studies. All 10 loci fulfilled this criterion as the number of alleles per locus ranged from 14 to 34. The mean number of alleles per locus varies between 6.4 and 16.3, with an overall value for all breeds of 10.98 ± 3.48 alleles per locus (Table 2). The microsatellites are polymorphic for each locus.

3.2 Genetic variation within and among breeds

Mean observed and expected heterozygosity (H_o and H_e) are above 0.5 for all breeds and loci, i.e. rather high (Tables 2 and 3). The smallest mean expected and observed heterozygosity for all loci is found in *brt* (0.588 and 0.617 respectively). The highest H_e is detected in *gyr* (0.813) and highest H_o in *lsm* (0.790). *Gyr* was found to be little differentiated (F_{st} of 0.02) and *fehr* and *fekr* are moderately differentiated (F_{st} of 0.078 and 0.112 respectively). Generally the Zackel breeds show relatively little genetic differentiation compared to the mean F_{st} value of 0.097. On average, roughly 10% of the total genetic variation is distributed between groups, with the rest among individuals within groups. According to Hartl and Clark (2007), moderate differentiation is reflected in an F_{st} from 0.05 to 0.15. Based on the allelic richness (mean number of alleles per breed), the least diverse population is also *brt* (5.3), but here the most diverse is *mm* (16.1). The mean number of alleles per breed depends on the number of individuals examined per breed.

Deviations from the HWE are significant ($P < 0.001$) for several loci in different populations (Table 4). Only five breeds (*bab*, *brt*, *cha*, *tcig*, and *tex*) are in HWE for all 10 loci, while *cig* deviates from HWE at all loci. Within the remaining 12 breeds, the number of markers that show a significant deviation from HWE equilibrium ranges from one to five. The loci 4 and 8 show the highest frequency of deviations from HWE, in 11 and 7 breeds respectively. While we impute the allelic state in the presence of null alleles, we cannot eliminate the possibility that null alleles still influence our estimates of deviation from HWE. Undetected population subdivision, i.e. the Wahlund effect (Hartl and Clark, 2007), may also explain deviation from HWE.

3.3 Pairwise genetic distances

All pairwise F_{st} comparisons are significant at $P < 0.05$ (Table 4). Results obtained with the co-ancestry coefficient (Reynolds' distance) are similar and therefore not shown. Pairwise F_{st} values range from 0.021 between the two Merino breeds (*nhm/mm*) and between Hungarian Merino and Transylvanian Zackel (*mm/gyr*) to 0.212 between British Milksheep and Awassi sheep (*brt/awa*). The *brt* population shows the greatest divergence ($F_{st} > 0.12$) to all other breeds, and *awa*, *cha*, and *tex* have divergences greater than 0.1 to several breeds. Low distance values of $F_{st} < 0.05$ were found especially between *bab*, *cig*, *gyr*, and *mm* and other breeds. The genetic distance between the three Zackel breeds is small, with pairwise F_{st} values of 0.03 (*fehr/fekr*), 0.037 (*fehr/gyr*) and 0.058 (*fekr/gyr*). The breeds furthest from Zackel breeds are *brt*, *cha*, and *awa*. The black and white Zackel tend to differ more from the other Hungarian breeds than their Transylvanian relatives, likely due to the low differentiation of the latter.

The pairwise genetic distance matrix (Table 4) is translated into a dendrogram (Fig. 1) using the UPGMA method. The British Milksheep and Awassi sheep branch first and are represented as outgroups to all other breeds. The two lowland Hungarian Zackel breeds (*fehr/fekr*) cluster closely together, while the Transylvanian Zackel *gyr* clusters with other sheep breeds.

3.4 Admixture model

Breeds that deviated from HWE were tested separately for population substructure. According to the method in Pritchard et al. (2000), we could not find evidence for additional population structure among these groups. Thus there seems to be no indication of the Wahlund effect in our data.

3.5 The USEPOPINFO model

The USEPOPINFO model assigns individuals to their most likely breed, assuming that their prior assignment to a breed is usually correct. Compared to the ADMIXTURE model,

Table 3. Genetic diversity and information to HWE for all 18 breeds. Number of individuals and mean number of alleles per breed are given. Ho (observed heterozygosity), He (expected heterozygosity), and single Fst for each breed. Deviations from HWE were significant ($P < 0.001$) for listed loci. Locus information see Table 2.

Breed	Number of individuals	Mean number of alleles	Ho	He	Deviation from HWE at locus No.	Fst
<i>awa</i>	38	8.3	0.726	0.754	8	0.104
<i>bab</i>	129	10.2	0.758	0.772	–	0.068
<i>brt</i>	30	5.3	0.617	0.588	–	0.274
<i>cha</i>	150	8.6	0.704	0.699	–	0.134
<i>cig</i>	543	15.8	0.772	0.804	all	0.027
<i>cik</i>	146	9.9	0.778	0.765	4	0.118
<i>fehr</i>	403	11.9	0.773	0.808	3, 4, 7, 9, 10	0.078
<i>fekr</i>	330	11.2	0.735	0.760	4, 7, 9	0.112
<i>gyr</i>	275	15.6	0.786	0.813	1, 4, 5	0.020
<i>ile</i>	454	12.1	0.754	0.769	1, 2, 4, 5, 10	0.077
<i>lac</i>	245	10.7	0.761	0.771	4, 9	0.075
<i>lsm</i>	168	11.3	0.790	0.791	4, 9	0.085
<i>mm</i>	907	16.1	0.782	0.801	1, 4, 6, 7, 10	0.030
<i>nff</i>	261	10.0	0.709	0.727	6, 9	0.097
<i>nhm</i>	670	12.2	0.756	0.779	2, 4, 8, 9	0.096
<i>suf</i>	528	12.8	0.706	0.725	1, 4, 8	0.075
<i>tcig</i>	89	9.9	0.787	0.799	–	0.058
<i>tex</i>	68	6.6	0.704	0.670	–	0.215
Mean	302	11	0.744	0.755		0.097

Table 4. Pairwise Fst values for all 18 breeds. All pairwise Fst's were significant at $P < 0.005$ after 110 permutations.

Breed	<i>awa</i>	<i>bab</i>	<i>brt</i>	<i>cha</i>	<i>cig</i>	<i>cik</i>	<i>fehr</i>	<i>fekr</i>	<i>gyr</i>	<i>ile</i>	<i>lac</i>	<i>lsm</i>	<i>mm</i>	<i>nff</i>	<i>nhm</i>	<i>suf</i>	<i>tcig</i>	<i>tex</i>	
<i>awa</i>	0.000																		
<i>bab</i>	0.095	0.000																	
<i>brt</i>	0.212	0.105	0.000																
<i>cha</i>	0.123	0.056	0.154	0.000															
<i>cig</i>	0.079	0.033	0.119	0.076	0.000														
<i>cik</i>	0.105	0.069	0.145	0.106	0.060	0.000													
<i>fehr</i>	0.082	0.057	0.133	0.091	0.051	0.078	0.000												
<i>fekr</i>	0.095	0.082	0.146	0.108	0.074	0.107	0.030	0.000											
<i>gyr</i>	0.082	0.030	0.125	0.077	0.031	0.059	0.037	0.058	0.000										
<i>ile</i>	0.121	0.045	0.166	0.117	0.061	0.089	0.075	0.107	0.040	0.000									
<i>lac</i>	0.100	0.035	0.121	0.073	0.024	0.069	0.061	0.090	0.032	0.065	0.000								
<i>lsm</i>	0.105	0.081	0.139	0.118	0.072	0.082	0.076	0.099	0.068	0.090	0.067	0.000							
<i>mm</i>	0.098	0.047	0.143	0.099	0.033	0.066	0.056	0.078	0.021	0.041	0.041	0.059	0.000						
<i>nff</i>	0.104	0.043	0.133	0.076	0.042	0.086	0.068	0.091	0.054	0.084	0.051	0.094	0.069	0.000					
<i>nhm</i>	0.132	0.078	0.180	0.133	0.056	0.096	0.078	0.107	0.048	0.053	0.059	0.084	0.021	0.099	0.000				
<i>suf</i>	0.099	0.039	0.137	0.084	0.050	0.086	0.078	0.111	0.060	0.092	0.066	0.099	0.070	0.052	0.110	0.000			
<i>tcig</i>	0.070	0.038	0.128	0.084	0.028	0.065	0.051	0.076	0.026	0.065	0.043	0.079	0.043	0.064	0.070	0.053	0.000		
<i>tex</i>	0.174	0.073	0.143	0.149	0.069	0.105	0.093	0.141	0.068	0.089	0.071	0.122	0.077	0.096	0.099	0.111	0.088	0.000	

the USEPOPINFO model is more stable because it uses information of predefined groups. The posterior probability of membership of each individual to its predefined breed is high: 95.9 to 99.0 %, depending on the breed. Also, only a few individuals per breed are not assigned correctly (Fig. 2). If individuals are not assigned to the correct breed, *structure* usually cannot clearly assign them to any of the other 17 breeds in our sample. Some of the questionable individuals may have mixed ancestry.

Cig, *fekr*, *gyr*, and *nhm* are the groups with the highest amount of wrongly assigned individuals (3.6, 3.0, 3.7, and 4.1 % respectively). Of those four breeds, only *cig* and *gyr* show small differentiation levels ($F_{st} = 0.027$, 0.020 respectively), and none show particularly low numbers of individuals in the sample, which may make assignment difficult. On the other hand, *fekr* and *nhm* are just as differentiated as other breeds, whose membership proportion is higher (Table 3 for F_{st} values). Hence there is no connection between assignment proportion and level of differentiation.

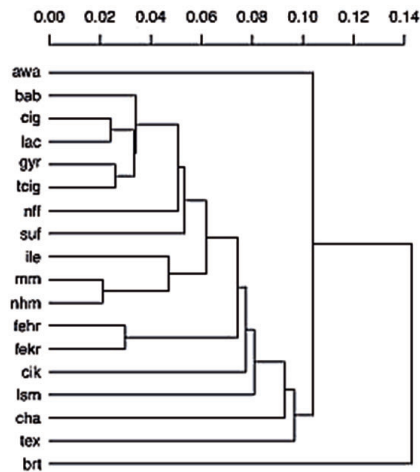


Figure 1. UPGMA dendrogram of Hungarian sheep breeds using the pairwise F_{st} distances as dissimilarity matrix. The scale represents F_{st} distance.

With the Zackel breeds, wrong assignments are generally confined to within the three breeds: in the *fekr* group, just two individuals are not clearly assigned to *fekr*. In the *fehr* group 13 individuals are not assigned to the *fehr* breed; 4 of these are assigned to *fekr* ($q > 0.701$), while the other 9 show no clear preference for any of the 18 breeds. In the *gyr* group, two individuals are assigned to *mm* with $q = 0.731$ and $q = 0.837$ respectively; the other five individuals are not clearly assigned to any breed. Of the Zackel individuals not clearly assigned, 13 might be hybrids between either two Zackel breeds or a Zackel and another breed. The genetic memberships of those sheep range from $q = 0.209$ to $q = 0.495$ in the different breeds. Only three Zackel individuals have probabilities less than 15% to be a pure Zackel or hybrid, or have Zackel ancestry.

3.6 USEPOPINFO PopFlag

With the PopFlag model in *structure* (Pritchard et al., 2000) we assigned 23 Zackel individuals to the 18 breeds. The memberships to our known breeds varies between $q = 0.173$ and $q = 0.905$, where $q > 0.5$ is taken as a clear assignment to a particular breed. Under this assumption, 20 individuals are assigned successfully to one breed. The three remaining sheep with q values lower than 0.500 might be hybrids among the other Zackel or further breeds. The sheep are primarily assigned to white or black Zackel, and some to Merino (see Table 5). All other breeds only receive limited support.

To understand the relatively high involvement of *mm* in this group of 23 unknown sheep tested with PopFlag, we checked genetic distances between the Zackel breeds and Hungarian Merino. Only *gyr* shows a slightly closer genetic distance (F_{st} value 0.021) to *mm*, but *fehr* and *fekr* are not more closely related to *mm* (pairwise F_{st} was 0.056 and 0.078

Table 5. Exact list of q values for each of the 23 individuals regarding their most probable breed. Individuals marked with an asterisk had only a membership of $q < 0.500$ in any breeds.

Sequential number and earmark number	<i>fehr</i>	<i>fekr</i>	<i>mm</i>
5435 AOv121848	–	0.552	0.192
5436 AOv121849	–	0.905	–
5437 AOv121850*	0.450	0.114	–
5438 AOv121852	0.695	–	–
5439 AOv121853	0.776	–	–
5440 AOv121854	–	0.903	–
5441 AOv121855	–	0.687	0.173
5442 AOv121856	0.231	0.516	–
5443 AOv121857*	0.306	0.351	–
5444 AOv121858	–	0.658	–
5445 AOv121859	0.674	–	0.194
5446 AOv121860	–	0.783	–
5447 AOv121861	–	0.572	0.251
5448 AOv121862	0.623	0.164	–
5449 AOv121863	–	0.846	–
5450 AOv121864	0.291	0.507	–
5451 AOv121865	0.713	–	–
5452 AOv121866*	–	0.446	0.293
5453 AOv121867	0.839	–	–
5454 AOv121868	–	0.819	–
5455 AOv121869	0.854	–	–
5456 AOv121870	–	0.893	–
5457 AOv121871	0.599	–	–

respectively) than to other breeds (see Table 4 with pairwise F_{st} values). The high proportion of Merino sheep may thus be explained by the origin (herd) of the sheep.

4 Discussion

The three Zackel breeds show high genetic variability in both the expected heterozygosities (gene diversity) (*fekr*: 0.760; *fehr*: 0.808; and *gyr*: 0.813) as well as the mean number of alleles (*fekr*: 11.2; *fehr*: 11.9; and *gyr*: 15.6). Their gene diversity is actually higher than the average 0.755 over all breeds, while their mean number of alleles is about average. We note that the number of alleles depends on the number of individuals screened and is therefore not easily comparable among breeds and studies. Among Spanish sheep analysed with 19 microsatellites, results showed 7.5 to 9.9 alleles per locus and the expected heterozygosity ranging from 0.773 to 0.814 (Arranz et al., 1998). Kusza et al. (2008) study Tsigai- and Zackel-type sheep with 16 microsatellite loci. In this study, gene diversities vary between 0.614 and 0.812 and mean allele numbers between 2.3 and 8.8. In Austrian sheep breeds analysed with 25 microsatellite loci, the expected heterozygosity ranges from 0.67 to 0.78 and the number of alleles from 6.19 to 10.71 (Baumung et al., 2006). While results

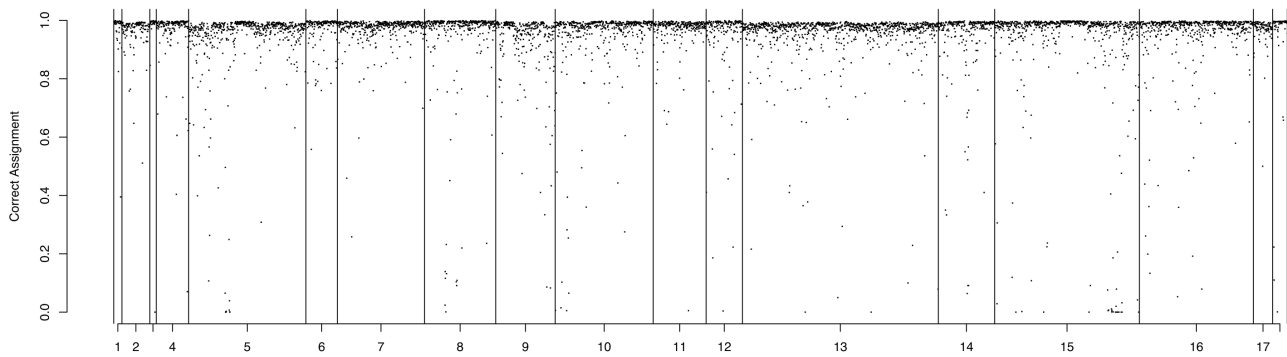


Figure 2. Assignment of individuals to their breeds using the USEPOPINFO model in *structure*. Each individual is represented by a dot of its posterior probability of correct assignment. Low values indicate migrants or hybrids.

from different microsatellite screens are generally difficult to compare, the numbers are rather similar among these studies.

The three Zackel breeds show average to high genetic variability compared to other Hungarian sheep breeds or breeds from other countries, e.g. Spain and Austria. Some of the sheep breeds used in those studies are endangered, such that low genetic variability would be expected. The generally low differentiation of the Zackel breeds indicates high effective population sizes. This is especially true for the Transylvanian Zackel, which has low census population sizes in Hungary. The main distribution range of this breed is, however, Transylvania, which today is located in Romania. The Transylvanian Zackel is also the least differentiated breed in our sample. In contrast, the British Milksheep is the most differentiated. It was established in Great Britain and then exported to Hungary and other countries. From our study we cannot deduce whether the British populations of the British Milksheep are already differentiated or whether the Hungarian populations show the effects of a later population bottleneck.

With the Hungarian breeds, 10 % of the genetic variability is among breeds and 90 % among individuals within breeds. Other studies on rare and endangered breeds, such as that of Kusza et al. (2008), show with multilocus F_{st} values that about 16 % of the total genetic variation in Tsigai and Zackel breed types in different European regions was among breeds. Among Austrian breeds (Baumung et al., 2006), only 8 % of genetic diversity segregates among breeds. But comparison with those studies is difficult because of our focus on relatively common breeds and comparatively high sample sizes.

Population distances based on pairwise F_{st} and Reynolds' distance show close relationships between the three Zackel breeds. This is expected for the white and black Zackel, because of their common history, although now they are bred strictly separately. But the Transylvanian Zackel should show larger distances because of its different origin from the Hungarian Zackel breeds. Interestingly, the Transylvanian Zackel has a closer relationship to some other breeds (*bab*, *cig*, *lac*, *mm*, and *tcig*) than to *fehr* or *fekr*. In Austrian studies (Bau-

mung et al., 2006; Schwend, 2001), the Texel breed is always the furthest from other breeds. In our study, this is not apparent. Rather, the British Milksheep is most differentiated from other breeds.

With the UPGMA method, a dendrogram (Fig. 1) was produced from the pairwise genetic distance matrix (Table 4). The British Milksheep and Awassi sheep branch first, the two lowland Hungarian Zackel breeds (*fehr* / *fekr*) cluster closely together, and the Transylvanian Zackel *gyr* clusters with other sheep breeds. Note that representation in a dendrogram assumes a bifurcating tree of relationships among the breeds. The population structure of sheep breeds in our study is, however, characterised by frequent hybridisation among breeds, at least during the phase of establishment of a breed, followed by a phase of maintenance of the breed.

Compared to the expected heterozygosities, the observed heterozygosities of our Zackel breeds show a slight deficit (Table 3). In line with this result, deviations from HWE for some loci could be detected: *fehr* deviated at five loci, and *fekr* and *gyr* at three. This deviation is in the range of that observed for the other breeds. Generally, all microsatellite loci deviate from HWE in at least some breeds. In particular, the loci OarAE129 and INTRA23 deviated in many breeds, which likely is caused by null alleles. Another explanation for Hardy–Weinberg disequilibrium would have been undetected population subdivision within our breeds, i.e. the Wahlund effect (Hartl and Clark, 2007). But with the methods in *structure*, we could not find evidence for this phenomenon.

With the microsatellite markers and the Bayesian clustering method (Pritchard et al., 2000) differentiation among and assignment to the closely related Hungarian sheep breeds is generally successfully resolved (Fig. 2). Only a few individuals cannot be unequivocally assigned to their nominal breed, while the vast majority of individuals are assigned correctly. In particular, correct assignments are very high for the Zackel breeds: 98.9 % of *fehr*, 97.0 % of *fekr*, and 96.3 % of *gyr*. For the rest, hybridisation could not be excluded. Those results also show that the USEPOPINFO model works very

well with the 18 populations. In our opinion, the underlying population genetic model employed in *structure* (Pritchard et al., 2000) quite accurately reflects the breeding history of the populations involved in this study.

5 Conclusions

Sheep breeds in Hungary are generally well differentiated, as the posterior assignments are generally to the correct breed. The Hungarian population of the British Milksheep is most differentiated from all other sheep breeds. The Awassi sheep is genetically slightly closer to all other breeds. The two Hungarian Zackel types, white and black, are genetically closely related and have high genetic variability. Nevertheless, there is little indication of hybridisation of these two closely related breeds to any other sheep breed. Thus, the breeding programmes successfully keep these two breeds distinct from other breeds while nevertheless maintaining a high genetic variability for these relatively rare breeds. Interestingly, the Transylvanian Zackel is not more closely related to the Hungarian Zackel types than to other breeds. This indicates a different breeding history, such that the Hungarian lowland and Transylvanian highland Zackel breeds do not seem to share much more than the name. Moreover, the Transylvanian Zackel shows the highest expected heterozygosity and the least differentiation in our sample group, probably due to fewer or less severe population bottlenecks than in other breeds. Hence, although just a small number of Transylvanian Zackel remained in Hungary, the breed has maintained a high genetic variability.

Acknowledgements. This work is part of the master's thesis of the first author. The Ethical and Animal Welfare Committee of the University of Veterinary Medicine, Vienna, was notified of the study.

Edited by: A.-E. Freifrau von Tiele-Winckler

Reviewed by: two anonymous referees

References

- Arranz, J. J., Bayon, Y., and San Primitivo, F.: Genetic relationships among Spanish sheep using microsatellites, *Anim. Genet.*, 29, 435–440, 1998.
- Baumung, R., Cubric-Curik, V., Schwend, K., Achmann, R., and Solkner, J.: Genetic characterisation and breed assignment in Austrian sheep breeds using microsatellite marker information, *J. Anim. Breed. Genet.*, 123, 265–271, 2006.
- Berger, B. and Fischerleitner, F.: Die seltenen, erhaltungswürdigen Schafrassen in Österreich [Rare and endangered sheep breeds in Austria], *ÖNGENE*, 15–16, available at: <http://www.oengene.at/images/stories/neu/downloads/7249-schafrassen.pdf> (last access: 20 August 2015), 2009.
- Buchanan, F. C., Adams, L. J., Littlejohn, R. P., Maddox, J. F., and Crawford, A. M.: Determination of Evolutionary Relationships among Sheep Breeds Using Microsatellites, *Genomics*, 22, 397–403, 1994.
- DAD-IS: Domestic Animal Diversity Information System (DAD-IS), Food and Agriculture Organization of the United Nations, <http://www.fao.org/dad-is>, last access: 20 August 2015.
- Excoffier, L., Laval, G., and Schneider, S.: Arlequin (version 3.0): An integrated software package for population genetics data analysis, *Evol. Bioinform.*, 1, 47–50, 2005.
- FAO: Measurement of Domestic Animal Diversity (MoDAD): Original Working Group Report, in: *Book Measurement of Domestic Animal Diversity (MoDAD): Original Working Group Report*, Rome, FAO, 1998.
- FAO: Molecular genetic characterization of animal genetic resources, *FAO Animal Production and Health Guidelines*, FAO, Rome, 85 pp., 2011.
- Guo, S. W. and Thompson, E. A.: Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles, *Biometrics*, 48, 361–372, 1992.
- Hartl, D. L. and Clark, A. G.: The Wahlund Principle, in: *Principles of population genetics*, Sinauer, Sunderland, USA, 288–293, 2007.
- Hortobágyi Nonprofit Kft.: Gyimes racka sheep, <http://www.hortobagy.eu/hu/regi-magyar-haziallataink/juhok/gyimesi-racka/> (last access: 20 August 2015), 2014.
- Hungarian Zackel sheep breeder association, Magyar Rackajuhtenyésztő Egyesület, <http://mjkszh.hu/tenyeszto/magyar-rackajuhtenyeszto-egyesulet>, last access: 20 August 2015.
- Kusza, S., Nagy, I., Sasvari, Z., Stigel, A., Nemeth, T., Molnar, A., Kume, K., Bosze, Z., Javor, A., and Kukovics, S.: Genetic diversity and population structure of Tsigai and Zackel type of sheep breeds in the Central-, Eastern- and Southern-European regions, *Small Ruminant Res.*, 78, 13–23, 2008.
- MacHugh, D. E., Loftus, R. T., Bradley, D. G., Sharp, P. M., and Cunningham, P.: Microsatellite DNA Variation within and among European Cattle Breeds, *P. Roy. Soc. B-Biol. Sci.*, 256, 25–31, 1994.
- Oklahoma State University: Breeds of Livestock Project, <http://www.ansi.okstate.edu/breeds/sheep/> (last access: 20 August 2015), 1995.
- Pritchard, J. K., Stephens, M., and Donnelly, P.: Inference of population structure using multilocus genotype data, *Genetics*, 155, 945–959, 2000.
- Pritchard, J. K., Wen, X., and Falush, D.: Documentation for *structure* software: Version 2.3., 38, available at: <http://pritchardlab.stanford.edu/structure.html> (last access: 20 August 2015), 2010.
- Reynolds, J., Weir, B. S., and Cockerham, C. C.: Estimation of the Co-Ancestry Coefficient – Basis for a Short-Term Genetic-Distance, *Genetics*, 105, 767–779, 1983.
- Sambras, H. H. *Farbatlas Nutztierassen: 263 Rassen in Wort und Bild*, Stuttgart, Germany, Ulmer Verlag, 2011.
- Schwend, K.: Genetic variability of the Carinthian sheep in Austria, *Inaugural-Disseration at University of Veterinary Medicine, Vienna*, 2001.