Historical sketch of Slovak Haban (Hutterite) population based on autosomal STR analysis

M. Soták[1], E. Petrejčíková[1], D. Siváková[2], K. Rębała[3], A. Bôžiková[1], J. Bernasovská[1], J.Čarnogurská[1], I. Boroňová[1], S. Mačeková[1], L. Homoľová[1], A. Sovičová[1], D. Gabriková[4], L. Rusínová[1], I. Bernasovský[1]

[1] *Department of Biology, Faculty of Humanities and Natural Science, University of Prešov, Slovakia*

[2] *Department of Anthropology, Faculty of Natural Science, Comenius University in Bratislava, Slovakia*

[3] *Department of Forensic Medicine, Medical University of Gdansk, Poland*

[4] *Centre of Animal and Human Ecology, University of Prešov*

Abstract

According to the Hutterite chronicles, the Habans arrived from Austrian Tyrol, Switzerland and northernmost Italy and stayed in four regions of Slovakia (Sobotište, Veľké Leváre, Moravský Svätý Ján, Trenčín). There are some communities in western Slovakia, which retained their Haban cultural identity and still identify themselves as descendents of the Hutterite population with their own specific customs. Slovak Habans are typical founder population with significant social isolation for which high degree of inbreeding is typical. Present study investigated STR polymorphisms as a powerful genetic tool for population genetic studies. The aim was to perform a comparative, population genetic study based on 15 STR loci widely used in forensic genetics, of the Haban population, the Slovak majority population and the population of Tyrol. We analyzed allele frequencies and other statistical parameters in three selected populations in order to identify groups of specific ethnic origin and establish their genetic relationship. The dataset included 110 unrelated Habans and 201 unrelated individuals from the Slovak majority population, as well as allelic frequencies for the population of Austrian Tyrol available in the literature. Population pairwise F_{ST} values used as a short term genetic distance between populations showed significant differentiation between the Habans and both reference populations ($F_{ST} = 0.0025$ and 0.0042 for comparison with the Slovaks and Austrians, respectively; $P < 10^{-3}$). The Slovak Hutterites were demonstrated to be genetically distinct and more closely related to their geographic neighbors than to their historical ancestral population, which may be at least partially explained by gene flow between neighboring Haban and Slovak populations.

Introduction

The Habans are a group of Hutterites, who converted to Catholicism in 19th century and until now retained a separate ethnic identity in western Slovakia. The Hutterites are descendants of Anabaptists from Switzerland, the western Tyrol region of Austria and northernmost Italy (Weisensse and Siváková 2003). Because of their rejection of both the Roman Catholic Church and the more moderate Protestant movement teachings, they suffered from religious persecution. This was the reason why they had to move to the other countries across Central Europe, then to Russia and finally to Northern America. At the beginning of the 16th century some Hutterites escaped to Slovakia (part of the Austrian empire) for the first time (Hutterites 2006). From 1526 to 1700 the Hutterites lived in western Slovakia where they built several settlements, with the first one established in Sobotište (Weisensse and Siváková 2003). During the subsequent re-catholicization of the Slovak population, in the late 1800's, many Hutterites moved to Transylvania (present-day Romania) and to the Molotschna district north of the Black Sea (present-day Ukraine) (Kalesný 1981) (Figure I.). After Russia introduced new regulations, many Hutterites moved to the USA in 1874. The first few years were hard for all the communities of Hutterites, but gradually they prospered and established many more colonies (Hutterites 2006). Currently, there are over 50,000 people in the USA and Canada who define themselves as Hutterites.

Figure I. Hutterite migration in Europe in the 16th and 18th centuries. The regions of analyzed samples were highlighted.
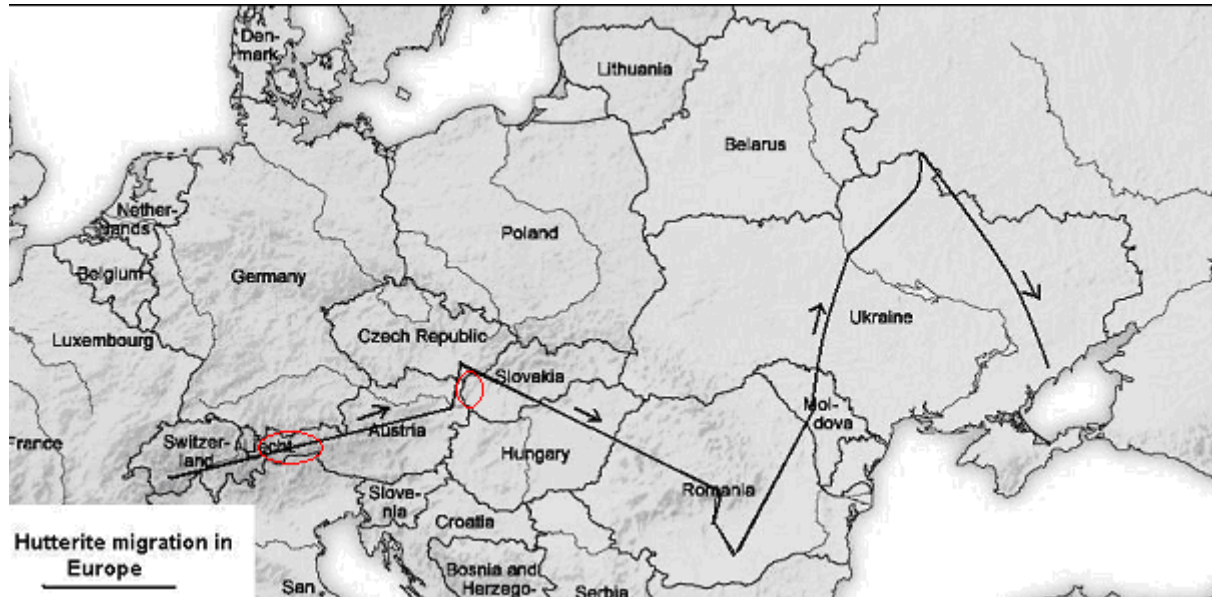
The isolated Hutterite populations are of special interest to anthropologists and human geneticists. Biological distances between individual populations based on screening of dermatoglyphical features assessed by means of the Hiernaux distance coefficient showed a clear separation of the Habans from the non-Habans inhabiting the same villages as well as from the other Slovak groups (Siváková and Pospíšil 2001). The explosion of molecular biology techniques enabled deeper insight into the Hutterites' genetic history by analysis of uniparentally inherited markers linked to the Y chromosome and mitochondrial DNA. Analysis of Y-chromosomal haplotypes defined by 12 STR loci in the Slovak Hutterite descendants revealed genetic distinctiveness of the Habans from the general western Slovak population, and their genetic proximity to populations inhabiting regions where they originated (Petrejčíková et al. 2009, 2011). A recent study of 9 microsatellite loci and 11 unique event polymorphisms of the Y chromosome and the hypervariable region I of the mitochondrial DNA in the North American Hutterites showed their unique genetic

background related to a similar extent to central and eastern European populations. An admixture analysis indicated a relatively high genetic contribution of central European populations to the Hutterite gene pool (Pichler et al. 2010). However, no data have been available for the autosomal markers commonly used in forensic practice.

Materials and Methods

Blood samples were collected from 311 unrelated healthy individuals from three western Slovak villages: Moravský Svätý Ján (48.586839° N, 16.9955444° E), Sobotište (48.7408015° N, 17.408881° E) and Veľké Leváre (48.503181° N, 17.0016977° E), including 110 Habans and 201 representatives of the Slovak majority population as a reference sample. Written informed consent was obtained from all participants. Analyzed samples were chosen from individuals of Haban ancestry according to chronicles which are recognized sources describing the familial structure of the Haban community.

Genomic DNA was extracted from all collected blood samples with the use of Jet Quick® (Genomed) and forensicGEM® (ZyGEM) kits. The amplification was performed on a T-personal 48 thermal cycler (Biometra) and 2720 Thermal Cycler® (Applied Biosystems) using PowerPlex® 16 System kit (Promega). The samples were purified by EXO-SAP solution and desalted by Microspin® G-50 columns (GE Healthcare). Capillary electrophoresis systems MegaBACE® 1000 genetic analyzer (GE Healthcare) and 3130 Genetic Analyzer® (Applied Biosystems) were used to separate amplified DNA products in accordance with the manufacturers' recommended protocols. Data analysis and allele detection were performed by MegaBACE® Genetic Profiler® software and GeneMapper® Software v4.0. Allele designation followed recommendations of the DNA commission of the International Society for Forensic Genetics (ISFG) (Bar et al. 1997) and considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature (Gill et al. 1997).

The simple gene counting method was used to calculate allele frequencies. The genetic data analysis was performed by the use of PowerStats®V12 (Promega) spreadsheet (Tereba 1999) and Arlequin® 3.5.1.2 (Excoffier et al. 2005). Observed (HO) and expected

heterozygosity (HE) and P-values for Hardy–Weinberg equilibrium for each single locus were determined by the exact test (probability based on 100,000 steps in the Markov chain, significance level=0.05). Matching probability (MP), power of discrimination (PD) (Nei 1987), polymorphic information content (PIC), power of exclusion (PE), typical paternity index (TPI) (Brenner and Morris 1990) and average gene diversity (Guo and Thompson 1992) were calculated. The population differentiation analysis between the Slovak Haban and reference majority populations was conducted (Markov chain length of 10,000 steps).

Both Haban and Slovak majority populations were compared with data available for the population of Tyrol (the Innsbruck area in Austria), where Habans/Hutterites originated (Ross et al. 2001, Steinlechner et al. 2001, Steinlechner et al. 2002). Population pairwise F_{ST} and corresponding P values were computed separately for 15 autosomal STR markers by conventional F-statistics from allelic frequencies (significance level = 0.05) with the use of the Arlequin software. Since the Arlequin treats missing alleles in allelic frequency data as separate alleles, all the missing alleles were removed prior to calculations. Weighted average F_{ST} values over loci were calculated as proposed by Weir and Cockerham (1984). Fisher's (1954, section 21.1) method of combining probabilities was used to obtain a single overall P value for 15 independent markers. A one-dimensional scaling analysis was conducted on a matrix of linearized (Slatkin 1995) weighted average F_{ST} values with the use of STATISTICA 9.1 software (StatSoft).

Results and discussion

In the present study, we analyzed and compared 15 widely used forensic autosomal STR loci in the Haban and Slovak majority populations from the same region (Tables I. and II.), and found several significant differences (Table III.). The allele frequencies of 4 loci showed differences between both studied populations and all the loci were highly polymorphic. The value of average gene diversity over loci was 0.7615 in the Haban population and 0.784966 in the reference majority one. The allelic size range for all of studied loci was 7.933 and 8.6 in the Haban and Slovak majority populations, respectively (s.d.= 0.471). The number of alleles was lower in the Haban population (8.667) than in the majority population (9.8, s.d.= 0.801), partially due to the sample size of both datasets. The exact test of sample differentiation did not reveal non-random distribution of alleles in analyzed population samples (P = 0.18525). No significant deviation from the Hardy-Weinberg equilibrium was found for all the markers tested in both populations, except for Penta E (P < 0.003 after Bonferroni's correction). The expected heterozygosity was higher than the observed one among the Habans (HE = 0.7925; HO = 0.7529), while both values were comparable in the Slovak majority population (HE = 0.7965; HO = 0.7733). The higher amount of genetic variability was found in the Slovak majority population. Small isolated populations with higher degree of inbreeding are characterized by excess of homozygotes in autosomal markers. Indeed, 11 out of 15 analyzed loci showed higher homozygosity in the Haban population than in the Slovak majority one.

Population pairwise F_{ST} values used as a short term genetic distance between populations showed significant differentiation between the studied populations. Comparison of our population samples with the Austrian population from Tyrol showed genetic distinctiveness of the Haban population from both Tyrolean and Slovak populations (Figure II.) in spite of the fact that Tyrol and western Slovakia are Habans' past and present

8

homelands. Out of the two reference populations, the Habans were demonstrated to be genetically more closely related to the Slovaks than to the Austrians (weighted average F_{ST} values of 0.0025 and 0.0042, respectively). Among all the pairwise comparisons, the Slovaks and Austrians revealed the closest genetic proximity (weighted average $F_{ST} = 0.0018$).

The autosomal STR genetic databases established for the studied populations have a potential application in forensic practice. The calculated results showed that the allele frequencies and other population genetic statistical data of descendants of the Habans were different from the gene pool of the Slovak and Austrian populations and that the Habans have remained genetically isolated from the Slovak majority population. Our results point to a founder effect, rather than a gene pool inherited from the ancestral population, as a likely reason for the Habans' genetic uniqueness. The founder effect hypothesis among the Habans is supported by an unexpectedly high frequency of Y chromosomes predicted as belonging to haplogroup E1b1b (43.6%) among the Haban males (Petrejčíková et. al. 2009), which generally does not exceed 10% in central European populations (Semino et al. 2004). We have also found out that the Slovak Hutterites are more closely related to the Slovaks than to the Austrians from Tyrol, which may be at least partially explained by gene flow between neighboring Haban and Slovak populations.

Table I. Allele frequencies and statistical parameters of 15 STR loci in the Slovak Haban population (No. of gene copies=220)

Allele	D3S1358	TH01	D21S11	D18S51	Penta E	D5S818	D13S317	D7S820	D16S539	CSF1PO	Penta D	VWA	D8S1179	TPOX	FGA
5					0,096										
6		0,266													
7		0,15			0,146			0,009							
8		0,117			0,015		0,155	0,141	0,041		0,01		0,014	0,509	
9		0,215		0,01	0,027	0,068	0,195	0,055	0,055		0,186				0,118
9,3		0,248													
10		0,005			0,167	0,023	0,041	0,259	0,055	0,202	0,119		0,115	0,05	
11				0,019	0,086	0,332	0,309	0,227	0,312	0,243	0,262		0,133	0,323	
12				0,131	0,126	0,4	0,25	0,145	0,335	0,399	0,224		0,161		
13				0,098	0,101	0,177	0,15	0,023	0,193	0,092	0,148	0,009	0,28		
14	0,142			0,145	0,121	0,032	0,027				0,043	0,136	0,188		
15	0,216			0,154	0,076							0,136	0,106		
15,2															
16	0,289			0,154	0,02							0,141			
17	0,225			0,131	0,02							0,314			
18	0,124			0,093								0,155			
19	0,005			0,056								0,105			0,061
20															0,159
21															0,15
22															0,159
22,2															0,009
23															0,21
24															0,131
25															0,079
26															0,014
27			0,014												0,009
28			0,185												

29		0,255													
30		0,236													
30,2		0,051													
31		0,069													
31,2		0,074													
32		0,019													
32,2		0,056													
33,2		0,037													
HO	0.76147	0,76636	0,81481	0,76636	0,75758	0.70000	0.80000	0.75455	0.72477	0.63303	0.83810	0.83636	0.77982	0.49091	0.86916
HE	0.78751	0,78917	0,83105	0,87938	0,89089	0.69934	0.79211	0.80506	0.74908	0.73276	0.81276	0.81337	0.82214	0.62304	0.85933
P	0.01468	0.03539	0.18083	0.29477	0.00084	0.66289	0.93002	0.66692	0.31895	0.15369	0.01182	0.65831	0.52265	0.01818	0.62623
MP	0,09	0,094	0,061	0,032	0,032	0,151	0,082	0,07	0,113	0,116	0,083	0,068	0,065	0,197	0,044
PD	0,91	0,906	0,939	0,968	0,968	0,849	0,918	0,93	0,887	0,884	0,92	0,932	0,935	0,803	0,956
PIC	0,75	0,75	0,81	0,86	0,88	0,64	0,76	0,77	0,7	0,69	0,780	0,78	0,79	0,55	0,84
PE	0,53	0,538	0,627	0,538	0,523	0,428	0,599	0,518	0,468	0,332	0,672	0,668	0,562	0,18	0,733
TPI	2,1	2,14	2,7	2,14	2,06	1,67	2,5	2,04	1,82	1,36	3,09	3,06	2,27	0,98	3,82

HO: observed heterozygosity; HE: expected heterozygosity; P: P-value of exact test for deviation from HW equilibrium; MP: matching

probability; PD power of discrimination; PIC: polymorphic information content; PE: power of exclusion; TPI: typical paternity index

Table II. Allele frequencies and statistical parameters of 15 STR loci in the Slovak majority population (No. of gene copies=402)

Allele	D3S1358	TH01	D21S11	D18S51	Penta E	D5S818	D13S317	D7S820	D16S539	CSF1PO	Penta D	VWA	D8S1179	TPOX	FGA
5					0,1										
6		0,22													
7		0,151			0,152			0,018							
8		0,096			0,019		0,168	0,15	0,009		0,014		0,027	0,628	
9		0,216			0,038	0,064	0,077	0,136	0,105	0,027	0,152		0,009	0,069	
9,3		0,307													
10		0,009			0,119	0,086	0,059	0,259	0,05	0,332	0,19		0,073	0,083	
11		0,009		0,014	0,076	0,345	0,3	0,223	0,318	0,259	0,152		0,082	0,193	
12				0,078	0,148	0,323	0,255	0,173	0,35	0,318	0,233	0,01	0,164	0,023	
13	0,032			0,119	0,095	0,159	0,1	0,036	0,159	0,05	0,233	0,005	0,323		
14	0,145			0,165	0,029	0,009	0,036		0,009	0,005	0,014	0,144	0,177		
15	0,2			0,165	0,067					0,009	0,01	0,111	0,118		
15,2															
16	0,223			0,161	0,057							0,202	0,018		
17	0,186			0,128	0,052							0,25			
18	0,205			0,083	0,033							0,202			
19	0,009			0,041	0,014							0,072			0,073
20				0,032											0,128
21															0,188
22															0,174
22,2															0,028
23															0,106
24															0,197
25															0,069
26															0,028
27			0,032												
28			0,193												
29			0,225												
30			0,174												
30,2			0,073												

31		0,069													
31,2		0,06													
32		0,032													
32,2		0,11													
33,2		0,018													
HO	0.75124	0.80000	0.84000	0.83000	0.77551	0.72637	0.76119	0.85572	0.72139	0.73632	0.71429	0.85641	0.80100	0.58500	0.84500
HE	0.80181	0.77599	0.85496	0.87119	0.90394	0.72748	0.80773	0.80877	0.74634	0.72873	0.82658	0.81478	0.80714	0.60555	0.86609
P	0.14420	0.15778	0.05326	0.25055	0.00004	0.09072	0.04648	0.02568	0.47736	0.19296	0.00573	0.13149	0.04283	0.79845	0.61334
MP	0,068	0,097	0,046	0,04	0,025	0,12	0,077	0,079	0,107	0,138	0,068	0,085	0,064	0,229	0,044
PD	0,932	0,903	0,954	0,96	0,975	0,88	0,923	0,921	0,893	0,862	0,93	0,915	0,936	0,771	0,956
PIC	0,78	0,74	0,84	0,86	0,89	0,7	0,77	0,78	0,69	0,67	0,780	0,79	0,79	0,52	0,84
PE	0,566	0,562	0,701	0,719	0,547	0,518	0,549	0,758	0,472	0,442	0,365	0,863	0,566	0,245	0,648
TPI	2,29	2,27	3,41	3,63	2,19	2,04	2,2	4,23	1,83	1,72	1,46	7,43	2,29	1,14	2,87

HO: observed heterozygosity; HE: expected heterozygosity; P: P-value of exact test for deviation from HW equilibrium; MP: matching

probability; PD power of discrimination; PIC: polymorphic information content; PE: power of exclusion; TPI: typical paternity index

Table III. F_{ST} and corresponding P values for pairwise comparisons between the Haban, Slovak and Austrian/Tyrolean populations (P values < 0.05 are bold).

Locus	Habans vs. Slovaks		Habans vs. Austrians		Slovaks vs. Austrians		Austrian population reference
	F_{ST}	P	FST	P	FST	P	
CSF1PO	0.0093	**0.0133**	0.0205	**0.0003**	0.0051	**0.0355**	R01
D3S1358	-0.0002	0.4343	0.0032	0.1023	0.0002	0.3656	S01
D5S818	0.0010	0.2591	0.0004	0.3269	-0.0020	0.8790	R01
D7S820	0.0002	0.3784	-0.0008	0.5384	-0.0013	0.7368	R01
D8S1179	0.0027	0.1135	0.0050	**0.0369**	0.0022	0.0939	S01
D13S317	-0.0012	0.6471	0.0005	0.3305	0.0021	0.1215	R01
D16S539	-0.0012	0.5899	0.0002	0.3489	-0.0009	0.7006	R01, S01
D18S51	-0.0002	0.4549	-0.0018	0.8536	0.0033	**0.0226**	S01
D21S11	0.0013	0.2178	-0.0009	0.6024	0.0006	0.2691	S01
FGA	0.0053	**0.0167**	0.0004	0.3485	0.0006	0.2728	S01
Penta D	0.0073	**0.0126**	0.0123	**0.0005**	0.0043	**0.0124**	S02
Penta E	0.0022	0.1115	0.0065	**0.0026**	0.0012	0.1216	S02
TH01	0.0019	0.1867	0.0060	**0.0210**	0.0019	0.0943	R01, S01
TPOX	0.0092	**0.0261**	0.0003	0.3165	0.0073	**0.0240**	R01
vWA	0.0017	0.1931	0.0112	**0.0010**	0.0035	**0.0184**	R01, S01
Weighted average F_{ST}	0.0025	0.0007	0.0042	$8 \cdot 10^{-8}$	0.0018	0.0002	

R01 Ross et al. 2001; S01 Steinlechner et al. 2001; S02 Steinlechner et al. 2002

Figure II. One-dimensional scaling plot of weighted average pairwise F_{ST} distances over 15 autosomal STR markers for the compared populations.
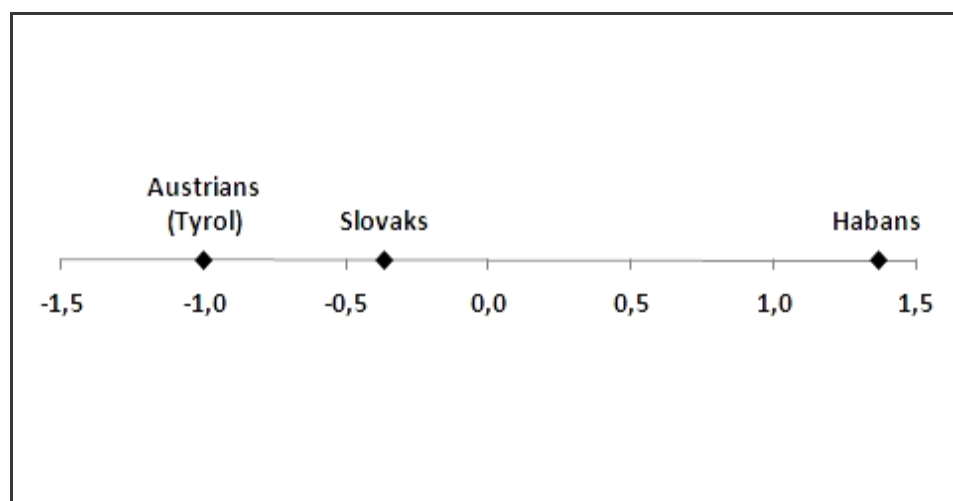
Quality control

Laboratory internal control standards and kit controls, QC test of Y-STR haplotype reference database (YHRD).

Access to the data

Access to data may be requested from soty@centrum.sk

References

Bar, W., Brinkmann, B., Budowle, B., et al. 1997. DNA recommendations further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. International Society for Forensic Haemogenetics. Int. J. Legal Med 110:175–176.

Brenner, CH., Morris, JW. 1990. Paternity index calculations in single locus hypervariable DNA probes: validation and other studies, in: Proceedings for the International Symposium on Human Identification 1989. Promega Corporation.pp. 21–53.

Excoffier, LG., Laval, S., Schneider, P. 2005. Arlequin ver. 3.0, an integrated software package for population genetics data analysis. Evol. Bioinformatics Online 1:47–50.
http://cmpg.unibe.ch/software/arlequin3/.

Fisher, RA, 1954. Statistical methods for research workers. 12th edit. Oliver & Boyd, Edinburgh. 356 p.

Hutterites. 2006. Hutterites Brethern Schmiedeleut Conference web site.
http://www.hutterites.org.

Gill, P., Brinkmann, B., d'Aloja, E., et al. 1997. Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Sci. Int. 87:185–192.

Guo, S., Thompson, E. 1992. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48:361–372.

Kalesný, F. 1981. Habans in Slovakia. Tatran Press Bratislava (in Slovak).

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press. New York, NY, USA.

Petrejčíková, E., Siváková, D., Soták, M., et al. 2009. Y-STR Haplotypes and Predicted Haplogroups in the Slovak Haban Population. Journal of Genetic Genealogy. 5/2:69-74.

Petrejčíková, E., Siváková, D., Soták, M., et al. 2011. Comparison of Y-STR polymorphisms in three different Slovak population groups. Anthrop. Anz. 68/2:111–127.

Pichler, I., Fuchsberger, C., Platzer, C., et al. 2010. Drawing the history of the Hutterite population on a genetic landscape: inference from Y-chromosome and mtDNA genotypes. European Journal of Human Genetics 18(4):463–470.

Ross, J., Parson, W., Furać, I., et al. 2001. Multiplex PCR amplification of eight STR loci in Austrian and Croatian Caucasian populations. Int J Legal Med 115:57–60.

Semino, O., Magri, C., Benuzzi, G., et al. 2004. Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. Am. J. Hum. Genet. 74:1023-1034.

Siváková, D., Pospíšil, M. 2001. Dermatoglyphic analysis of Habans (Hutterites) from Slovakia. Anthropologischer Anzeiger 59: 355–363.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462.

Steinlechner, M., Berger, B., Scheithauer, R. et al. 2001. Population Genetics of ten STR loci (AmpFlSTR SGM plus) in Austria. Int. J. Legal Med 114:288–290.

Steinlechner, M., Grubwieser, P., Scheithauer, R., et al. 2002. STR loci Penta D and Penta E: Austrian Caucasian population data Int. J. Legal Med 116: 174–175.

Tereba, A. 1999. Tools for Analysis of Population Statistics. Profiles in DNA 3. Promega

Corporation.http://www.promega.com/geneticidtools/powerstats/

Weir, BS., Cockerham, CC. 1984. Estimating F-statistics for the analysis of population structure. Evolution

38:1358–1370.

Weisensse, KE., Siváková, D. 2003. A Comparison of Slovakian Haban Populations Using Finger Ridge-Count

Data. Stud Tribes Tribals 1:73–76.

76.