Genetic Variation at 15 Polymorphic, Autosomal, Short Tandem Repeat Loci of Two Hungarian Populations in Transylvania, Romania

Szabolcs József Demeter[1], Beatrice Kelemen[1], Gyöngyi Székely[2], Octavian Popescu[1,3]

[1]Molecular Biology Center, Interdisciplinary Research Institute on Bio-Nano-Sciences, Babeş-Bolyai University, Cluj-Napoca, Romania

[2]Experimental Biology Department, Faculty of Biology and Geology, Babeş-Bolyai University, Cluj-Napoca, Romania

[3]Institute of Biology, The Romanian Academy, Bucharest, Romania

Aim To determine allele distribution and genetic parameters for two populations living in the Romanian region of Transylvania: Hungarians from Cluj and Szeklers from Covasna county, and to compare the results between the two populations and with other Hungarian and Romanian populations.

Methods Allele frequencies for 15 autosomal STR loci (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, VWA, TPOX, D18S51, D5S818, and FGA), several forensic parameters, and paternity parameters were determined for Szekler Hungarians of Covasna county (CV-Sze, n = 278) and non-Szekler Transylvanian Hungarians, who were represented by Hungarians from Cluj county (CJ-Hu, n = 146).

Results Average expected heterozygosity was above 70%. The combined power of discrimination and combined power of exclusion values were high. All tested loci were in agreement with Hardy-Weinberg equilibrium, with the exception of the CSF1PO locus for Covasna county. Pairwise population comparison tests and exact population differentiation tests showed no significant differences between the CJ-Hu and CV-Sze populations, and the CV-Sze group showed greater differences from other Romanian populations than did the CJ-Hu group.

Conclusion Hungarians from Cluj show greater genetic heterogeneity than Szeklers from Covasna. The loci tested are suitable for studying micro-differentiation between these two populations, and between these populations and other populations in Hungary and Romania.

Since its introduction 25 years ago, DNA fingerprint analysis has become a major tool in diagnosing and treating disease, forensic identification, taxonomy, phylogenetics, and other applications (1,2). Microsatellites, also called simple sequence repeats or short tandem repeats (STR), are among the most polymorphic DNA markers. These sequences of 2-6 basepairs are easily amplified by polymerase chain reaction (PCR) and show widespread and uniform distribution throughout the genome. They show a high level of polymorphism, which is relatively stable (3). These properties of STR loci make them suitable for numerous genetic, forensic, and medical applications.

According to the 2002 census in Romania (4), 19.6% of the residents of the Transylvania region belong to the Hungarian ethnic group, with most of them living in the Szekler (Székely) counties of Covasna and Harghita. Earlier genetic studies have suggested that ancient Hungarians and Szekler Hungarians separated from each other more than 1000 years ago. Some centuries ago, the two Szekler groups separated from each other, which affected the genetic structure of these groups (5). In the last decade, genetic parameters for the Szekler population (HR-Sze) and Csángó population (HR-Csn) from Harghita have been published (6), but no population study has been conducted on Szekler communities from Covasna and other, non-Szekler Hungarians, from Transylvania.

This study sheds light on the genetic makeup of Hungarian communities from Transylvania by determining CODIS STR allele frequencies, as well as forensic and paternity data for the Szekler Hungarians of Covasna county (CV-Sze) and non-Szekler Hungarians of Cluj county (CJ-Hu).

Using allele frequencies, we carried out pairwise comparisons and differentiation tests to test the following hypotheses:

1. There is significant genetic distance between the non-Szekler Hungarians (CJ-Hu) and the Szekler populations (HR-Sze and CV-Sze), which indicates genetic isolation of the two Szekler groups from the other Hungarian communities living in Transylvania.

2. There is significant genetic distance between the two Szekler populations (HR-Sze and CV-Sze), which reveals genetic isolation of the Szekler communities.

3. There is a positive correlation between geographical distance and genetic distance when ethnic Hungar-

ian populations in this study are compared with other ethnic Romanian populations.

4. As a result of population migration and ethnic cross-breeding, non-Szekler Hungarians in this study show greater genetic heterogeneity than do Szekler Hungarians.

MATERIALS AND METHODS

Population

The samples for this study were collected from 424 unrelated, healthy Caucasian individuals of Hungarian ethnicity from Covasna (n = 278) and Cluj (n = 146) counties, in Transylvania, Romania. The sample sizes for both counties were sufficiently large for forensic study (7). All participants were volunteers who gave their oral consent in compliance with the ethical norms set by Romanian legislation. The study protocol was approved by the ethics committee of Babeş-Bolyai University.

Extraction

Genomic DNA was extracted from whole blood using the Promega SV Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA), following the manufacturer's recommendations (8). Standards of the Interdisciplinary Research Institute on Bio-Nano-Sciences were strictly applied when handling the blood samples, and standard precautions for DNA analysis (9,10) were taken.

Polymerase chain reaction

PCR amplifications were performed using the AmpF*l*STR Identifiler PCR Amplification kit (Applied Biosystems, Carlsbad, CA, USA) according to the manufacturer's recommendations (11).

Typing

PCR products were analyzed using an ABI PRISM® 310 Genetic Analyzer (Applied Biosystems). The following 15 loci were tested: D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, VWA, TPOX, D18S51, D5S818, and FGA. Alleles of these loci were identified according to recommendations of the International Society for Forensic Genetics (9), the International Society of Forensic Hemogenetics (10), the European Network of Forensic Science Institutes, and the Council of the Europe-

an Union (12,13), with the aid of the AmpF*l*STR Identifiler Allelic Ladder.

Data analysis

Allele frequencies and statistical data on heterozygosity, matching probability, power of discrimination (PD), power of exclusion (PE), polymorphism information content, and typical paternity index (PI) were obtained using the PowerStats workbook template, version 1.2 (Promega) in Microsoft Excel (14). Tests measuring the possible departures from the Hardy-Weinberg expectations were performed with the Arlequin software, version 3.5 (15).

In order to match the two populations with each other and with other Hungarian and Romanian populations, pairwise comparison tests (F_{st}) and exact population differentiation tests (exact tests) were performed using the Arlequin software. Previously published allele frequencies from the following populations were compared with CV-Sze and CJ-Hu: Szeklers from Harghita county (HR-Sze) (6), Csángós from Harghita county (HR-Csn) (6), Romanians from the Bucharest area (B-Ro) (16), Western Romanians (West) (17), and the general populations of Transylvania (TRS) (18), Moldavia (Moldavia) (19), the Dobruja region, located between the lower Danube River and the Black Sea (Dobruja) (20), and the Wallachia region of Romania (Wallachia) (21).

RESULTS

Allele distribution and forensic and paternity-related genetic parameters were determined for two Hungarian populations from Transylvania (Figure 1): Szeklers from Covasna county (n = 278) and non-Szekler Hungarians from Cluj county (n = 146). Allele frequencies, Hardy-Weinberg equilibrium (HWE) test results, and statistical parameters such as matching probability, PD, polymorphism information content, PE, and PI are summarized in Tables 1 and 2.

The average expected heterozygosity (Hexp) value in CJ-Hu population was lower than the observed heterozygosity (Hexp = 0.7857; Hobs = 0.7941). The opposite was observed in the CV-Sze population (Hexp = 0.7965; Hobs = 0.7890).

All loci met the conditions of HWE, with the exception of CSF1PO and D16S539 in CV-Sze and CSF1PO and D18S51 in CJ-Hu. However, after applying the Bonferroni correction, there was no departure from HWE at any of the loci studied, with the exception of CSF1PO in CV-Sze. As a re-

Figure 1.

Historical regions of Romania: Transylvania, Moldavia, Wallachia, and Dobruja; CJ – Cluj county; CV – Covasna county; HR – Harghita county.

sult, there was no reason for HWE rejection (21). The average PI values obtained for CV-Sze and CJ-Hu were 2.58 and 3.25, respectively. The PD and PE values for all the analyzed markers were high, with PD ranging from 0.8804 to 0.9717 and PE values from 0.3617 to 0.7501. The combined PD values for the 15 loci were 0.99999999999999992 for CJ-Hu and 0.999999999999999997 for CV-Sze. The combined PE values for the 15 loci were 0.9999998 for CJ-Hu and 0.999998 for CV-Sze. These combined PD values and combined PE values were high enough to make these loci suitable for identification of individuals.

The genetic structure of these two populations was compared with each other as well as with other Hungarian populations from Transylvania (HR-Sze and HR-Csn) and Romanian populations (B-Ro, Dobruja, Moldavia, Trs, Wallachia, and West) (Table 3). No significant differences were found between CJ-Hu and CV-Sze; both populations showed one significantly different locus (D5S818) when compared with Szeklers from Harghita county. Neither CJ-Hu nor CV-Sze showed differences from Romanians living in Bucharest. Significant differences between CV-Sze and Romanians from Wallachia were found in the allele frequencies at 3 loci (F_{st}: CSF1PO, TH01, and D16S539) and 5 loci (exact test: D21S11, CSF1PO, D16S539, D2S1338, and D18S51). The differences in allele frequency were greater between CV-Sze and the other populations (10 or 14 significantly different loci by the F_{st} test or exact test, respectively) than between CJ-Hu and the other populations (4 or 2 significantly different loci by the F_{st} test or exact test, respectively). Table 4 summarizes the significantly different loci found in these comparisons.

TABLE 1. Allele frequencies and related statistical data for the Szekler population of Covasna county (n = 278)

	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
Allele:															
6	–	–	–	–	–	0.2086	–	–	–	–	–	0.0018	–	–	–
7	–	–	0.0144	0.0018	–	0.1619	–	–	–	–	–	0.0036	0.0018	0.0018	–
8	0.0126	–	0.1529	0.0126	–	0.1007	0.1511	0.0198	–	–	–	0.4982	–	0.0018	–
9	0.0144	–	0.1763	0.0701	–	0.1979	0.0791	0.1565	–	–	–	0.1205	–	0.0431	–
9.3	–	–	–	–	–	0.3273	–	–	–	–	–	–	–	–	–
10	0.0827	–	0.2716	0.3004	–	0.0036	0.0576	0.0594	–	–	–	0.0594	0.0180	0.0863	–
10.2	–	–	–	–	–	–	–	–	–	–	–	–	0.0036	–	–
11	0.0755	–	0.2104	0.3040	–	–	0.3147	0.2481	–	0.0018	–	0.2913	0.0126	0.2752	–
12	0.1440	–	0.1421	0.2517	0.0018	–	0.2932	0.2985	–	0.0791	–	0.0252	0.1529	0.3904	–
12.2	–	–	–	–	–	–	–	–	–	0.0036	–	–	–	–	–
13	0.2931	–	0.0323	0.0468	–	–	0.0863	0.1817	–	0.2301	0.0018	–	0.1259	0.1906	–
13.2	–	–	–	–	–	–	–	–	–	0.0270	–	–	0.0018	–	–
14	0.2249	–	–	0.0090	0.0791	–	0.0180	0.0360	0.0018	0.3308	0.1169	–	0.1565	0.0036	–
14.2	–	–	–	–	–	–	–	–	–	0.0360	–	–	–	–	–
15	0.1295	–	–	0.0036	0.2464	–	–	–	0.0018	0.1511	0.1205	–	0.1241	0.0072	–
15.2	–	–	–	–	–	–	–	–	–	0.0450	–	–	–	–	–
16	0.0198	–	–	–	0.2985	–	–	–	0.0630	0.0504	0.2338	–	0.1474	–	–
16.2	–	–	–	–	–	–	–	–	–	0.0343	–	–	–	–	–
17	0.0035	–	–	–	0.2195	–	–	–	0.1834	0.0054	0.2428	–	0.0989	–	0.0054
17.2	–	–	–	–	–	–	–	–	–	0.0036	–	–	–	–	–
18	–	–	–	–	0.1457	–	–	–	0.1295	–	0.2050	–	0.0809	–	0.0108
18.2	–	–	–	–	–	–	–	–	–	0.0018	–	–	–	–	–
19	–	–	–	–	0.0090	–	–	–	0.0953	–	0.0702	–	0.0450	–	0.0486
20	–	–	–	–	–	–	–	–	0.1169	–	0.0090	–	0.0288	–	0.1546
21	–	–	–	–	–	–	–	–	0.0306	–	–	–	–	–	0.1601
22	–	–	–	–	–	–	–	–	0.0486	–	–	–	–	–	0.1690
22.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.0090
23	–	–	–	–	–	–	–	–	0.1295	–	–	–	0.0018	–	0.1367
23.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.0018
24	–	–	–	–	–	–	–	–	0.0845	–	–	–	–	–	0.1601
25	–	–	–	–	–	–	–	–	0.0971	–	–	–	–	–	0.0899
26	–	0.0072	–	–	–	–	–	–	0.0180	–	–	–	–	–	0.0468
27	–	0.0288	–	–	–	–	–	–	–	–	–	–	–	–	0.0072
28	–	0.1331	–	–	–	–	–	–	–	–	–	–	–	–	–
29	–	0.2266	–	–	–	–	–	–	–	–	–	–	–	–	–
29.2	–	0.0054	–	–	–	–	–	–	–	–	–	–	–	–	–
30	–	0.2068	–	–	–	–	–	–	–	–	–	–	–	–	–
30.2	–	0.0575	–	–	–	–	–	–	–	–	–	–	–	–	–
31	–	0.0630	–	–	–	–	–	–	–	–	–	–	–	–	–
31.2	–	0.0845	–	–	–	–	–	–	–	–	–	–	–	–	–
32	–	0.0162	–	–	–	–	–	–	–	–	–	–	–	–	–
32.2	–	0.1205	–	–	–	–	–	–	–	–	–	–	–	–	–
33.2	–	0.0414	–	–	–	–	–	–	–	–	–	–	–	–	–
34.2	–	0.0054	–	–	–	–	–	–	–	–	–	–	–	–	–
35.2	–	0.0018	–	–	–	–	–	–	–	–	–	–	–	–	–
38		0.0018													
Genetic parameters:															
observed heterozygosity	0.8309	0.8381	0.7842	0.7554	0.8058	0.7590	0.7518	0.7878	0.8777	0.7698	0.7950	0.6547	0.8777	0.6942	0.8525
expected heterozygosity	0.8127	0.8563	0.8061	0.7466	0.7744	0.7738	0.7748	0.7866	0.8859	0.8006	0.8111	0.6482	0.8795	0.7262	0.8646
P (Hardy-Weinberg equilibrium exact test)	0.2342	0.4394	0.1145	0.0014	0.1019	0.3677	0.2346	0.0049*	0.0738	0.0397	0.1182	0.1952	0.6350	0.5244	0.0637
matching probability	0.0640	0.0386	0.0693	0.1196	0.0998	0.0882	0.0858	0.0850	0.0281	0.0669	0.0677	0.1779	0.0290	0.1192	0.0369
power of discrimination	0.9360	0.9614	0.9307	0.8804	0.9002	0.9118	0.9142	0.9150	0.9719	0.9331	0.9323	0.8221	0.9710	0.8808	0.9631
polymorphism information content	0.7884	0.8407	0.7778	0.7032	0.7380	0.7384	0.7419	0.7545	0.8750	0.7762	0.7841	0.5935	0.8672	0.6816	0.8495
power of exclusion	0.6577	0.6716	0.5700	0.5190	0.6098	0.5253	0.5128	0.5765	0.7501	0.5442	0.5897	0.3617	0.7501	0.4194	0.6998
typical paternity index	2.96	3.09	2.32	2.04	2.57	2.07	2.01	2.36	4.09	2.17	2.44	1.45	4.09	1.64	3.39

*Bonferroni-type correction (0.0033).

TABLE 2. Allele frequencies and related statistical data for a non–Szekler Hungarian population represented by Hungarians from Cluj county (n = 146)

	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
Allele:															
6	–	–	–	–	–	0.2466	–	–	–	–	–	–	–	–	–
7	–	–	0.0171	–	–	0.1301	–	–	–	–	–	–	–	–	–
8	0.0205	–	0.1541	0.0068	–	0.1267	0.1267	0.0103	–	–	–	0.5480	–	0.0034	–
9	0.0068	–	0.1027	0.0274	–	0.1849	0.0856	0.0993	–	0.0034	–	0.0891	0.0034	0.0377	–
9.3	–	–	–	–	–	0.3014	–	–	–	–	–	–	–	–	–
10	0.0616	–	0.3356	0.2774	–	0.0103	0.0513	0.0719	–	–	–	0.0445	0.0137	0.0788	–
10.2	–	–	–	–	–	–	–	–	–	–	–	–	0.0034	–	–
11	0.0616	–	0.2398	0.3253	–	–	0.3391	0.2945	–	0.0172	–	0.3013	0.0103	0.3322	–
12	0.1747	–	0.1062	0.3049	–	–	0.2946	0.2877	–	0.0651	0.0034	0.0171	0.0822	0.3733	–
13	0.3356	–	0.0445	0.0514	–	–	0.0651	0.1986	–	0.2637	0.0034	–	0.1507	0.1610	
13.2	–	–	–	–	–	–	–	–	–	0.0342	–	–	–	–	–
14	0.2090	–	–	0.0068	0.1164	–	0.0308	0.0377	–	0.3253	0.1096	–	0.1953	0.0068	–
14.2	–	–	–	–	–	–	–	–	–	0.0342	–	–	–	–	–
15	0.1200	–	–	–	0.2568	–	0.0068	–	–	0.1165	0.1267	–	0.1541	0.0068	–
15.2	–	–	–	–	–	–	–	–	–	0.0548	–	–	–	–	–
16	0.0068	–	–	–	0.2329	–	–	–	0.0753	0.0548	0.2260	–	0.1267	–	–
16.2	–	–	–	–	–	–	–	–	–	0.0172	–	–	–	–	–
17	0.0034	–	–	–	0.2671	–	–	–	0.1883	0.0034	0.2364	–	0.1233	–	–
18	–	–	–	–	0.1164	–	–	–	0.0925	0.0034	0.2021	–	0.0548	–	0.0068
18.2	–	–	–	–	–	–	–	–	–	0.0068	–	–	–	–	–
19	–	–	–	–	0.0104	–	–	–	0.0788	–	0.0616	–	0.0480	–	0.0719
20	–	–	–	–	–	–	–	–	0.1575	–	0.0274	–	0.0102	–	0.1233
21	–	–	–	–	–	–	–	–	0.0240	–	–	–	0.0171	–	0.1507
22	–	–	–	–	–	–	–	–	0.0240	–	0.0034	–	0.0034	–	0.2193
22.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.0068
23	–	–	–	–	–	–	–	–	0.1370	–	–	–	–	–	0.1473
24	–	–	–	–	–	–	–	–	0.0925	–	–	–	0.0034	–	0.1678
24.2	–	0.0034	–	–	–	–	–	–	–	–	–	–	–	–	–
25	–	–	–	–	–	–	–	–	0.1130	–	–	–	–	–	0.0753
26	–	0.0034	–	–	–	–	–	–	0.0137	–	–	–	–	–	0.0274
27	–	0.0274	–	–	–	–	–	–	0.0034	–	–	–	–	–	–
28	–	0.1267	–	–	–	–	–	–	–	–	–	–	–	–	–
29	–	0.1815	–	–	–	–	–	–	–	–	–	–	–	–	0.0034
29.2	–	0.0034	–	–	–	–	–	–	–	–	–	–	–	–	–
30	–	0.2467	–	–	–	–	–	–	–	–	–	–	–	–	–
30.2	–	0.0377	–	–	–	–	–	–	–	–	–	–	–	–	–
31	–	0.0856	–	–	–	–	–	–	–	–	–	–	–	–	–
31.2	–	0.0993	–	–	–	–	–	–	–	–	–	–	–	–	–
32	–	0.0240	–	–	–	–	–	–	–	–	–	–	–	–	–
32.2	–	0.1267	–	–	–	–	–	–	–	–	–	–	–	–	–
33.2	–	0.0308	–	–	–	–	–	–	–	–	–	–	–	–	–
35	–	0.0034	–	–	–	–	–	–	–	–	–	–	–	–	–
Genetic parameters:															
observed heterozygosity	0.7603	0.9110	0.7945	0.6575	0.8288	0.7808	0.7808	0.8151	0.9384	0.7671	0.8493	0.5753	0.9315	0.6644	0.8562
expected heterozygosity	0.7907	0.8532	0.7820	0.7208	0.7812	0.7811	0.7671	0.7745	0.8778	0.7978	0.8202	0.5987	0.8714	0.7167	0.8525
P (Hardy-Weinberg equilibrium exact test)	0.2471	0.1511	0.0996	0.0210*	0.2993	0.0519	0.4589	0.1653	0.0627	0.0720	0.2215	0.5769	0.0036*	0.2844	0.2631
matching probability	0.0739	0.0494	0.0810	0.1287	0.0905	0.0855	0.0864	0.1014	0.0377	0.0680	0.0669	0.2198	0.0417	0.1207	0.0448
power of discrimination	0.9261	0.9506	0.9190	0.8713	0.9095	0.9145	0.9136	0.8986	0.9623	0.9320	0.9331	0.7802	0.9583	0.8793	0.9552
polymorphism information content	0.7627	0.8371	0.7513	0.6686	0.7461	0.7468	0.7339	0.7397	0.8655	0.7732	0.7956	0.5362	0.8579	0.6687	0.8351
power of exclusion	0.5275	0.8179	0.5889	0.3657	0.6535	0.5639	0.5639	0.6273	0.8742	0.5367	0.6935	0.2623	0.8601	0.3753	0.7070
typical paternity index	2.09	5.58	2.43	1.46	2.92	2.28	2.28	2.70	8.11	2.13	3.32	1.18	7.30	1.49	3.48

*Bonferroni-type correction (0.0033).

TABLE 3. Genetic comparison tests at 15 short tandem repeat loci of our CV-Sze and CJ-Hu groups with other populations in Hungary (HR-Csn, HR-Sze) and Romania (B-Ro, Dobruja, Moldavia, Trs, Wallachia, West) populations*

Marker	Test	Population data	CV-Sze n=278	HR-Sze n=257	HR-Csn n=220	B-Ro n=243	Trs n=1977	West n=219	Moldavia n=1321	Dobruja n=569	Wallachia n=1910
D8S1179	F_{st}	CV-Sze vs		0.485	0.252	0.197	0.567	0.838	0.798	0.833	0.516
	F_{st}	CJ-Hu vs	0.771	0.845	0.813	0.696	0.207	0.955	0.503	0.209	0.206
	Exact test	CV-Sze vs		0.607	0.238	0.498	0.582	0.990	0.961	0.901	0.677
	Exact test	CJ-Hu vs	0.911	0.547	0.516	0.622	0.198	0.984	0.672	0.305	0.213
D21S11	F_{st}	CV-Sze vs		0.612	0.104	0.730	0.863	0.455	0.747	0.833	0.634
	F_{st}	CJ-Hu vs	0.729	0.801	0.274	0.357	0.179	0.137	0.230	0.298	0.158
	Exact test	CV-Sze vs		0.934	0.313	0.601	0.308	0.756	0.344	0.817	0.018†
	Exact test	CJ-Hu vs	0.955	0.902	0.260	0.340	0.395	0.483	0.566	0.550	0.258
D7S820	F_{st}	CV-Sze vs		0.916	0.461	nd	0.487	0.828	0.614	0.732	0.290
	F_{st}	CJ-Hu vs	0.217	0.200	0.102	nd	0.055	0.451	0.180	0.064	0.076
	Exact test	CV-Sze vs		0.966	0.310	nd	0.641	0.810	0.879	0.886	0.734
	Exact test	CJ-Hu vs	0.387	0.271	0.047	nd	0.159	0.562	0.442	0.232	0.287
CSF1PO	F_{st}	CV-Sze vs		0.366	0.412	nd	0.017	0.294	0.035	0.202	0.040
	F_{st}	CJ-Hu vs	0.505	0.737	0.947	nd	0.934	0.945	0.896	0.963	0.985
	Exact test	CV-Sze vs		0.326	0.533	nd	<0.001	0.323	<0.001	0.047	0.003
	Exact test	CJ-Hu vs	0.658	0.833	0.931	nd	0.941	0.994	0.913	0.992	0.972
D3S1358	F_{st}	CV-Sze vs		0.886	0.024	0.160	0.067	0.870	0.474	0.763	0.312
	F_{st}	CJ-Hu vs	0.368	0.585	0.031	0.220	0.014	0.260	0.152	0.193	0.016
	Exact test	CV-Sze vs		0.870	0.006	0.167	0.214	0.897	0.627	0.876	0.402
	Exact test	CJ-Hu vs	0.496	0.765	0.024	0.285	0.058	0.348	0.231	0.472	0.065
TH01	F_{st}	CV-Sze vs		0.711	0.648	0.315	0.137	0.049	0.250	0.047	0.034
	F_{st}	CJ-Hu vs	0.746	0.855	0.841	0.875	0.835	0.741	0.965	0.626	0.712
	Exact test	CV-Sze vs		0.463	0.549	0.398	0.321	0.020	0.553	0.172	0.220
	Exact test	CJ-Hu vs	0.613	0.942	0.884	0.853	0.834	0.896	0.960	0.874	0.839
D13S317	F_{st}	CV-Sze vs		0.651	0.685	nd	0.718	0.360	0.850	0.917	0.578
	F_{st}	CJ-Hu vs	0.916	0.617	0.990	nd	0.260	0.592	0.435	0.508	0.199
	Exact test	CV-Sze vs		0.689	0.784	nd	0.919	0.508	0.920	0.979	0.908
	Exact test	CJ-Hu vs	0.856	0.753	0.954	nd	0.301	0.758	0.278	0.532	0.260
D16S539	F_{st}	CV-Sze vs		0.612	0.038	0.335	0.387	0.519	0.312	0.355	0.011
	F_{st}	CJ-Hu vs	0.497	0.916	0.538	0.322	0.701	0.980	0.837	0.850	0.476
	Exact test	CV-Sze vs		0.629	0.010	0.532	0.270	0.542	0.364	0.373	0.010
	Exact test	CJ-Hu vs	0.644	0.912	0.314	0.429	0.667	0.941	0.900	0.796	0.477
D2S1338	F_{st}	CV-Sze vs		0.347	0.660	0.127	0.088	0.573	0.163	0.404	0.146
	F_{st}	CJ-Hu vs	0.909	0.631	0.674	0.271	0.855	0.771	0.953	0.823	0.916
	Exact test	CV-Sze vs		0.350	0.656	0.065	0.001	0.554	0.016	0.245	0.016
	Exact test	CJ-Hu vs	0.968	0.679	0.781	0.305	0.719	0.733	0.894	0.815	0.844
D19S433	F_{st}	CV-Sze vs		0.456	0.316	0.653	0.049	0.457	0.613	0.103	0.171
	F_{st}	CJ-Hu vs	0.876	0.339	0.203	0.434	0.621	0.458	0.595	0.418	0.767
	Exact test	CV-Sze vs		0.561	0.429	0.749	0.135	0.604	0.663	0.154	0.221
	Exact test	CJ-Hu vs	0.902	0.184	0.291	0.691	0.272	0.567	0.365	0.135	0.143
vWA	F_{st}	CV-Sze vs		0.077	0.184	0.485	0.191	0.530	0.425	0.407	0.072
	F_{st}	CJ-Hu vs	0.996	0.220	0.480	0.725	0.464	0.696	0.699	0.625	0.277
	Exact test	CV-Sze vs		0.219	0.355	0.619	0.510	0.764	0.818	0.716	0.304
	Exact test	CJ-Hu vs	0.961	0.347	0.456	0.859	0.563	0.584	0.863	0.814	0.332
TPOX	F_{st}	CV-Sze vs		0.161	0.633	nd	0.996	0.740	0.708	0.849	0.963
	F_{st}	CJ-Hu vs	0.491	0.372	0.207	nd	0.420	0.827	0.777	0.292	0.348
	Exact test	CV-Sze vs		0.413	0.881	nd	0.605	0.930	0.656	0.877	0.692

TABLE 3. - continued. Genetic comparison tests at 15 short tandem repeat loci of our CV-Sze and CJ-Hu groups with other populations in Hungary (HR-Csn, HR-Sze) and Romania (B-Ro, Dobruja, Moldavia, Trs, Wallachia, West) populations*

Marker	Test	Population data	CV-Sze n=278	HR-Sze n=257	HR-Csn n=220	B-Ro n=243	Trs n=1977	West n=219	Moldavia n=1321	Dobruja n=569	Wallachia n=1910
	Exact test	CJ-Hu vs	0.824	0.315	0.486	nd	0.730	0.792	0.882	0.647	0.694
D18S51	F_{st}	CV-Sze vs		0.403	0.494	0.321	0.242	0.839	0.389	0.346	0.070
	F_{st}	CJ-Hu vs	0.287	0.850	0.488	0.209	0.688	0.384	0.727	0.449	0.705
	Exact test	CV-Sze vs		0.208	0.604	0.069	0.052	0.877	0.014	0.198	0.001
	Exact test	CJ-Hu vs	0.390	0.924	0.718	0.293	0.874	0.437	0.851	0.906	0.931
D5S818	F_{st}	CV-Sze vs		<0.001	0.959	nd	0.805	0.814	0.982	0.835	0.958
	F_{st}	CJ-Hu vs	0.610	<0.001	0.877	nd	0.262	0.830	0.696	0.271	0.353
	Exact test	CV-Sze vs		<0.001	0.828	nd	0.126	0.586	0.221	0.849	0.327
	Exact test	CJ-Hu vs	0.970	<0.001	0.895	nd	0.330	0.693	0.559	0.602	0.531
FGA	F_{st}	CV-Sze vs		0.909	0.229	0.208	0.500	0.364	0.706	0.731	0.424
	F_{st}	CJ-Hu vs	0.775	0.984	0.816	0.729	0.471	0.938	0.741	0.547	0.721
	Exact test	CV-Sze vs		0.960	0.184	0.061	0.256	0.374	0.751	0.598	0.413
	Exact test	CJ-Hu vs	0.948	1.000	0.878	0.539	0.959	0.918	0.994	0.964	0.987

*Abbreviations: B-Ro – Romanians from Bucharest area; CJ-Hu – non-Szekler Hungarians from Cluj county; CV-Sze – Szeklers from Covasna county; exact test – exact population differentiation test; Fst – pairwise population comparison test; HR-Csn – Csángós from Harghita county; HR-Sze – Szeklers from Harghita county; nd – not determined; Trs – population from Transylvanian region; West – population from Western Romania.
†Values in bold represent significant values at *P* < 0.05

TABLE 4. Summary of significantly different loci identified during population comparison tests*

Test performed	Population	CV-Sze n=278	HR-Sze n=257	HR-Csn n=220	B-Ro n=243	Trs n=1977	West n=219	Moldavia n=1321	Dobruja n=569	Wallachiaž n=1910	Total number of significantly different loci
F_{st}	CJ-Hu vs	0	D5S818	D3S1358	0	D3S1358	0	0	0	D3S1358	4
	CV-Sze vs		D5S818	D3S1358, D16S539	0	D19S433	TH01	CSF1PO	TH01	CSF1PO, TH01, D16S539	10
Exact test	CJ-Hu vs	0	D5S818	D7S820	0	0	0	0	0	0	2
	CV-Sze vs		D5S818	D3S1358, D16S539	0	CSF1PO, D2S1338	TH01	CSF1PO, D2S1338, D18S51	0	D21S11, CF1PO, D16S539, D2S1338, D18S51	14

*Abbreviations: B-Ro – Romanians from Bucharest area; CJ-Hu – non-Szekler Hungarians from Cluj county; CV-Sze – Szeklers from Covasna county; exact test – exact population differentiation test; F_{st} – pairwise population comparison test; HR-Csn – Csángós from Harghita county; HR-Sze – Szeklers from Harghita county; Trs – population data for Transylvanian region; West – population data for Western Romania.

DISCUSSION

Contrary to our hypothesis that there would be a significant genetic difference between non-Szekler and Szekler populations, our results showed no significant differences between CJ-Hu and Szekler populations (CV-Sze, HR-Sze), with only a single significantly different locus (D5S818). The results can be explained by high genetic heterogeneity of the non-Szekler Hungarian community of Cluj county (Hexp<Hobs). Further tests, with less heterogeneous non-Szekler Hungarian populations, are required to confirm or disprove the hypothesis 1.

When allele frequencies of the two Szekler populations were compared with each other, both pairwise compari-

son and exact differentiation tests revealed significant differences for one locus (D5S818). Although more significant loci were expected to be found, these results support our hypothesis that there is significant genetic distance between the two Szekler populations.

We hypothesized that there should be a positive correlation between genetic and geographical distance (22), but unexpected results were obtained when both populations in this study were compared with Romanian populations. Although we expected to find Dobruja to be genetically most remote from both populations included in the study, our results showed that Wallachia was genetically most remote. The differences in allele frequency were greater between CV-Sze and Wallachia (F_{st} test, 3 of

15 loci; exact test, 5 of 15 loci) than between CV-Sze and Dobruja (F_{st} test, 1 of 15 loci; exact test, 0 of 15 loci). In the same comparison, CJ-Hu showed significant differences at one locus (D3S1358) when compared with Wallachia and no significant differences when compared with Dobruja.

The allelic differences identified between CV-Sze and the other populations (F_{st} test, 10 significantly different loci; exact test, 14) were greater than the differences identified between CJ-Hu and the other populations (F_{st} test, 4 significantly different loci; exact test, 2), which supports our hypothesis that non-Szekler Hungarians would show greater genetic heterogeneity than Szekler Hungarians. In addition, the Hobs was lower than Hexp for CV-Sze, which may suggest inbreeding and isolation. This may explain why CV-Sze showed more significant differences from the other Romanian populations than CJ-Hu.

The results of this comparison may be affected by the number of loci compared; for example, data on only 10 loci were available for the B-Ro population.

Future investigations should include all 15 loci (also including Y-chromosome STRs) and should involve other Hungarian populations from Hungary, Serbia, Slovakia, Ukraine, and Austria. These studies will help clarify the genetic characteristics of the ethnic Hungarian community in Transylvania.

Acknowledgments

References

1 Decorte R. Genetic identification in the 21st century–Current status and future developments. Forensic Sci Int. 2010;201:160-4. Medline:20347537 doi:10.1016/j.forsciint.2010.02.029

2 Jeffreys AJ, Wilson V, Thein SL. Individual-specific 'fingerprints' of human DNA. Nature. 1985;316:76-9. Medline:2989708 doi:10.1038/316076a0

3 Tautz D. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res. 1989;17:6463-71. Medline:2780284 doi:10.1093/nar/17.16.6463

4 Romanian census 2002 [in Romanian]. Available from: http://www.recensamant.ro/. Accessed: October 29, 2010.

5 Beer Z, Zsolt Penzes Z, Csete K, Varga T. Comparison of two isolated "Hungarian" population to population of Budapest (mixed Hungarian) by Y-chromosomes. Int Congr Ser. 2003;1239:473-80. doi:10.1016/S0531-5131(02)00800-2

6 Egyed B, Furedi S, Padar Z. Population genetic study in two Transylvanian populations using forensically informative autosomal and Y-chromosomal STR markers. Forensic Sci Int. 2006;164:257-65. Medline:16314060 doi:10.1016/j.forsciint.2005.10.020

7 Chakraborty R. Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. Hum Biol. 1992;64:141-59. Medline:1559686

8 SV Wizard® Genomic DNA Purification Kit (Promega), Technical Manual No.050. Madison (WI): Promega; 2009.

9 Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, et al. DNA recommendations: Further report of the DNA Commission of the ISFG regarding the use of short tandem repeat systems. Forensic Sci Int. 1997;87:179-84.

10 Recommendations of the DNA Commission of the International Society for Forensic Haemogenetics relating to the use of PCR-based polymorphisms. Forensic Sci Int. 1992;55:1-3. Medline:1511935 doi:10.1016/0379-0738(92)90088-E

11 AmpFLSTR® Identifiler™ PCR Amplification Kit (Applied Biosystems), User Manual. Carlsbad (CA): Applied Biosystems; 2006.

12 Council of The European Union, No.doc: 15870/09 ENFOPOL 287 CRIMORG 170.

13 EU Council Resolution on the exchange of DNA analysis results; 2009/C 296/01.

14 Powerstats version 1.2. Available from: http://www.promega.com/geneticidtools/powerstats/. Accessed: October 29, 2010.

15 Excoffier L, Laval G, Schneider S. Arlequin ver. 3.5: an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online. 2005;47-50. Available from: http://cmpg.unibe.ch/software/arlequin35/Arl35Downloads.html. Accessed: October 29, 2010.

16 Barbarii LE, Rolf B, Constantinescu C, Hohoff C, Calistru P, Dermengiu D. Allele frequencies of 13 short tandem repeat (STR) loci in the Romanian population. Forensic Sci Int. 2004;141:171-4. Medline:15062958 doi:10.1016/j.forsciint.2003.11.005

17 Marian C, Anghel A, Bel SM, Ferencz BK, Ursoniu S, Dressler M, et al. STR data for the 15 AmpFlSTR identifiler loci in the Western Romanian population. Forensic Sci Int. 2007;170:73-5. Medline:16806768 doi:10.1016/j.forsciint.2006.04.011

18 Stanciu F, Stoian IM, Popescu OR. Comprehensive STR data for the AmpFlSTR Identifiler from Transylvania (NW Romania). Leg Med (Tokyo). 2009;11:48-50. Medline:19004660

19 Stanciu F, Popescu OR, Stoian IM. Allele frequencies of 15 STR loci in Moldavia region (NE Romania). Forensic Sci Int; Genet. 2009;4:e39-40. Medline:19948322 doi:10.1016/j.fsigen.2009.02.008

20 Stanciu F, Popescu OR, Stoian IM. STR data for the AmpFlSTR Identifiler from Dobruja region (SE Romania). Forensic Sci

Int; Genet. 2009;3:146-7. Medline:19215887 doi:10.1016/j.fsigen.2008.09.009

21   Stanciu F, Stoian IM, Popescu OR. Population data for 15 short tandem repeat loci from Wallachia Region, South Romania. Croat Med J. 2009;50:321-5. Medline:19480027 doi:10.3325/cmj.2009.50.321

22   Zhivotovsky LA, Feldman MW. Microsatellite variability and genetic distances. Natl Acad Sci. 1995;92:11549-52. doi:10.1073/pnas.92.25.11549