Int J Legal Med (1998) 112 : 72–74

SHORT COMMUNICATION

S. Füredi · Z. Kozma · J. Woller · Z. Pádár · M. Angyal
I. Bajnóczky · K. Nishi

Population genetic data on four STR loci in a Hungarian Romany population

Received: 15 January 1998 / Received in revised form: 3 April 1998

Abstract A population study of Hungarian Romanies was carried out on the STR loci HumLPL, HumF13B, HumFES and HumF13A01. There was little evidence for association of alleles within/between the four STR systems. Allele frequency distributions were significantly different between the Romany and the previously reported Central Hungarian population databases. Population differentiation was estimated by computing *F*- and *Φ*-statistics as well as frequency estimate differences of individual phenotypes for these two population samples. The results suggest that the population structure may have an effect on the interpretation of forensic DNA evidence in Hungary. Phylogenetic tree reconstruction with six populations from three major ethnic groups revealed a relatively distant genetic relationship of the Baranya Romanies with other Caucasian populations.

Key words Multiplex STR profiling · Automated DNA sequencer · Romany (Gypsy) population · Population genetics · Hungary

Introduction

With the advent of STR profiling it has become possible for forensic scientists to rapidly gain a reliable impression of the scale of population genetic effects caused by several factors such as inbreeding and substructuring [1]. Due to their relatively high proportion (approx. 6%) among the Hungarian inhabitants the Romanies represent one of the most relevant groups of the Hungarian population. Previous papers [2, 3] have shown that population studies in the Hungarian Romanies can be of great importance from the viewpoint of the examination of population differentiation. This paper provides additional population genetic data of a Hungarian Romany population on the STR loci HumLPL [4], HumF13B [5], HumFES [6] and HumF13A01 [7].

S. Füredi (✉) · J. Woller · Z. Pádár
Department of Biology, Institute for Forensic Sciences,
P.O. Box 314/4, H-1903 Budapest, Hungary

Z. Kozma · M. Angyal · I. Bajnóczky
Department of Forensic Medicine,
Medical University School of Pécs, P.O. Box 99,
H-7643 Pécs, Hungary

K. Nishi
Department of Legal Medicine,
Shiga University of Medical Science, Seta, Tsukinowa-cho,
Otsu, Shiga, 520–21 Japan

Materials and methods

Blood samples were collected from 135 unrelated Romany individuals residing in Baranya county (south-western Hungary). DNA samples (2–4 ng) were amplified using reagents provided in the GenePrint Singleplex and Fluorescent FFFL Quadruplex STR Systems (Promega, Madison, Wis.) according to the manufacturer's instructions. The PCR products were analysed on sequencing gels with silver staining [8] as well as on an ALF DNA sequencer (Pharmacia) [3]. The allele sizing accuracy of the ALF DNA sequencer was evaluated as previously described [3].

Possible divergence from Hardy-Weinberg expectations (HWE) was determined by the exact test [9]. An expectation maximization (EM) algorithm [10] was used to test allelic independence between STR loci. The frequency profile comparisons were performed using a computerized G-statistic test. Population substructure was measured by calculating the unbiased single-locus "coancestry coefficient" F_{ST} [11] and its *Φ*-statistic analogue Φ_{ST} [12, 13] using the software ARLEQUIN v1.0. Interpopulation differences were evaluated computing the differences between the respective log likelihood ratio estimates (log LRs) of each multi-locus profile from the databases to be compared. LRs for every member of the databases were considered as the inverse of the profile frequency estimates calculated from each population database. Phylogenetic trees were reconstructed using the software PHYLIP v3.5c [14] to evaluate genetic distances between population pairs. The standard genetic distance Ds [15] and the Ddm or $(\delta\mu)^2$ [16] distance were used as distance matrices for the construction of unrooted Neighbour-Joining trees.

Results and discussion

Both singleplex amplification with manual typing and multiplex fluorescent analysis gave the same typing re-

Table 1 Observed allele frequencies, homozygosities, and statistical data for the four STR loci in 135 unrelated Hungarian Romanies

Allele	LPL	F13B	FES	F13A01
3.2				0.081
4				0.044
5				0.381
6		0.207		0.233
7				0.196
8		0.241	0.007	
9	0.063	0.189		
10	0.289	0.359	0.182	
11	0.278	0.004[a]	0.437	
12	0.370		0.374	
13				0.004
14				0.030
15				0.026
17				0.004
Observed homozygosity	34.8%	23.7%	37.0%	25.2%
Exact test[b]	0.015	0.595	0.055	0.606
F_{ST}[c]	0.018	0.008	0.020	0.034
Φ_{ST}[c]	0.004	0.028	0.006	0.012

[a] Allelic designation of this PCR fragment was confirmed by sequencing analysis determining a previously published $[AAAT]_{11}$ sequence motif [21]
[b] These values are probability values
[c] Statistics for the estimation of the locus specific genetic variability between the Baranya Romany and the Central Hungarian populations [8, 17]

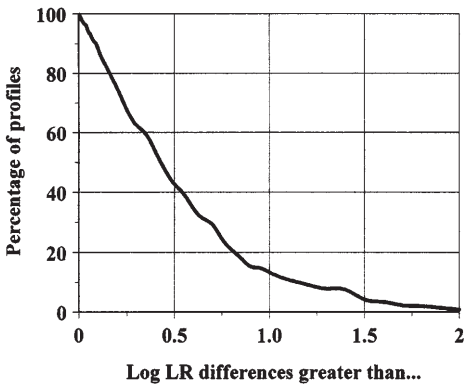

Fig. 1 Distribution of the log LR differences derived from the Baranya Romany and Central Hungarian population databases for 358 four-locus STR profiles. The percentage of samples exceeding any given value shown on the X axis is plotted

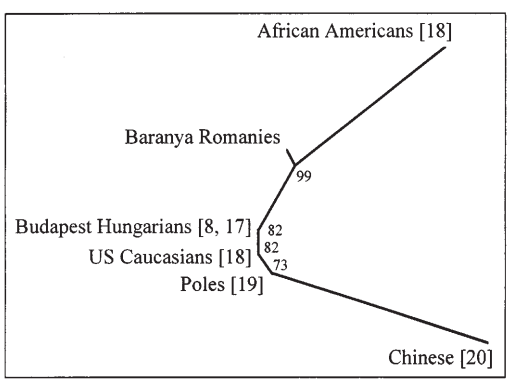

Fig. 2 Unrooted Neighbor-Joining tree based on the genetic distance Ds. The numbers at each node refer to the bootstrap values obtained in 1000 replications

sults. An extremely high allele sizing accuracy (99.97% on average for individual samples) was achieved by applying external (allelic ladder) and internal (150 and 271 bp) markers on the ALF DNA sequencer.

In the Romany population sample the exact test showed evidence of departure from HWE for the locus HumLPL (Table 1). Using the EM method there was evidence for the FES/F13A01 pair ($P = 0.003$) that the population sample deviates from expectations of allelic independence between the STR loci. These departures are likely be sampling effects but could also be due to inbreeding effects in such a genetically closed subpopulation.

The Hungarian Romany allele frequency values (Table 1) for all loci were significantly different from a Central Hungarian population database [8, 17] ($P \leq 10^{-3}$). The combined forensic efficiency values observed in the Romany population sample [PD(4 loci) = 0.9995; MEC(4 loci) = 0.9125] were slightly smaller than those found in the Central Hungarian database. Calculating Wright's *FST* indices for the two populations a relatively high level of *FST* values (Table 1) was found for the STR loci as compared to the previous observations in other Caucasian populations [1]. For the F13B and F13A01 loci this genetic correlation was reinforced by Φ-statistics, where genetic variation was considered at the molecular level. In the interpopulation comparison with respect to individual phenotypes 13.4% of the four-locus profiles had LR estimates differing by more than one order of magnitude (Fig. 1). These results suggest that the possibility of population differentiation should be taken into account in the calculation of match probabilities in Hungarian forensic cases.

In search of the relationship between the Baranya Romany population and the major population groups, phylogenetic trees were constructed. Using the allele frequency data of the four STR loci with an addition of the four supplementary microsatellites (VWA, TH01, TPOX, CSF1PO) analysed previously [3] both distance matrices gave the similar tree topology (Fig. 2). The US Caucasian [18], Polish [19] and Central Hungarian [8, 17] populations formed the Caucasian group to which the Romanies are genetically related. The Chinese [20] and the African American [18] samples were clearly separated from the Caucasian group by showing long branches.

Acknowledgements The authors wish to thank Prof. Dr. B. Brinkmann, Dr. B. Rolf and M. Schürenkamp (University of Münster, Germany) for their help with the sequencing analysis. Computer programs were kindly provided by B. Budowle (FBI Academy, Quantico, USA), G. Carmody (Carleton University, Ottawa, Canada) and L. Excoffier (University of Geneva, Switzerland).

References

1. Evett IW, Lambert JA, Buckleton JS, Weir BS (1996) Statistical analysis of a large file of data from STR profiles of British Caucasians to support forensic casework. Int J Legal Med 109:173–177
2. Hummel K, Kazarinowa-Fukshansky N (1991) Comparison of allele frequencies in 24 genetic markers for 12 Hungarian ethnic groups and for 8 other non-Hungarian reference populations. In: Czeizel A, Benkmann H-G, Goedde HW (eds) Genetics of the Hungarian population. Springer, Berlin Heidelberg New York, pp 281–306
3. Füredi S, Angyal M, Kozma Z, Sétáló J, Woller J, Pádár Z (1997) Semi-automatic DNA profiling in a Hungarian Romany population using the STR loci HumVWFA31, HumTH01, HumTPOX, and HumCSF1PO. Int J Legal Med 110:184–187
4. Zuliani G, Hobbs HH (1990) Tetranucleotide repeat polymorphism in the LPL gene. Nucleic Acids Res 18:4958
5. Nishimura DY, Murray JC (1992) A tetranucleotide repeat for the F13B locus. Nucleic Acids Res 20:1167
6. Polymeropoulos MH, Rath DS, Xiao H, Merril CR (1991) Tetranucleotide repeat polymorphism at the human c-fes/fps proto-oncogene. Nucleic Acids Res 19:4018
7. Polymeropoulos MH, Rath DS, Xiao H, Merril CR (1991) Tetranucleotide repeat polymorphism at the human coagulation factor XIII A subunit gene (F13A1). Nucleic Acids Res 19: 4306
8. Füredi S, Budowle B, Woller J, Pádár Z (1996) Hungarian population data on six STR loci – HUMVWFA31, HUMTH01, HUMCSF1PO, HUMFES/FPS, HUMTPOX, and HUMHPRTB – derived using multiplex PCR amplification and manual typing. Int J Legal Med 109:100–101
9. Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372
10. Slatkin M, Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. Heredity 76:377–383
11. Weir BS (1996) Genetic data analysis II. Sinauer Associates, Sunderland, Mass.
12. Excoffier L, Smouse P, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491
13. Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics 142:1061–1064
14. Felsenstein J (1989) Phylip – phylogeny inference package (version 3.2). Cladistics 5:164–166
15. Nei M (1972) Genetic distance between populations. Am Nat 106:283–291
16. Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldmann MW (1995) Genetic absolute dating based on microsatellites and the origin of modern humans. Proc Natl Acad Sci USA 92:6723–6727
17. Füredi S, Woller J, Pádár Z (1997) A population study of the STR loci HUMLPL, HUMF13B, and HUMF13A01 in Hungary. Int J Legal Med 110:107–108
18. Schumm JW, Lins AM, Micka KA, Sprecher CJ, Rabbach DR, Bacher JW (1997) Automated fluorescent detection of STR multiplexes – development of the GenePrint, PowerPlex and FFFL multiplexes for forensic and paternity applications. In: Proceedings from the First European Symposium on Human Identification, 1996. Promega Corporation, Madison, WI, pp 90–104
19. Miścicka-Śliwka D, Czarny J, Grzybowski T, Woźniak M (1998) Population genetics of 14 STRs: vWA, TH01, TPOX, CSF1PO, D5S818, D13S317, D7S820, D16S539, F13A01, FESFPS, F13B, LPL, D3S1358 and FGA in the Pomerania-Kujawy region of Poland. In: Olaisen B, Brinkmann B, Lincoln P (eds) Progress in forensic genetics 7. Elsevier, Amsterdam (in press)
20. Lee JC-I, Chen C-H, Tsai L-C, Linacre A, Chang J-G (1997) The screening of 13 short tandem repeat loci in the Chinese population. Forensic Sci Int 87:137–144
21. Alper B, Meyer E, Schürenkamp M, Brinkmann B (1995) HumFES/FPS and HumF13B: Turkish and German population data. Int J Legal Med 108:93–95