

Carrier Rates of Four Single-Gene Disorders in Croatian Bayash Roma

Ana Barešić and Marijana Peričić Salihović

To assess how specific population history, different migration routes, isolation, and endogamy practices contributed to the distribution of several rare diseases found in specific Roma groups, we conducted a population-based research study of rare disease mutations in Croatian Vlach Roma. We tested a total of 427 subjects from Baranja and Međimurje for the presence of four mutations causing hereditary motor and sensory neuropathy type Lom (HMSNL), GM1 gangliosidosis (GM1), congenital cataracts, facial dysmorphism and neuropathy (CCFDN), and limb girdle muscle dystrophy type 2C (LGMD2C), using the RFLP-PCR method to estimate carrier frequencies. We identified a total of four individuals heterozygous for the mutation causing HMSNL in the Baranja population, with a carrier rate amounting to 1.5%. Carriers for other three mutations causing GM1, CCFDN, and LGMD2C were not found in our sample. The carrier rate for the HMSNL mutation in Baranja is lower than in other Vlach Roma groups. In addition, distinct differences in carrier rates between the Croatian Vlach groups point to different genetic history, despite their belonging to the same Roma migration category and subgroup. The difference in carrier rates is either the result of admixture or the reflection of a greater extent of genetic drift since recent founding, maintained by a high degree of endogamy.

Introduction

ROMA POPULATIONS EXPRESS a range of rare monogenic diseases that only recently became recognized in medicine and at the moment, only several countries have registries of such diseases (Orpha.net), however, not for the ones affecting the Roma population. This represents a serious drawback since the Roma are especially prone to monogenic disorders as they are a young population founded by a small number of individuals and in addition, many different Roma groups practice endogamy and are isolated from majority of the populations. Data regarding rare monogenic diseases in the Roma populations living in Croatia do not exist, so we set out to conduct a population-based research study to assess the load of the Roma community with several rare diseases most common in specific groups. We used the obtained results to identify how the distribution of disease mutations was influenced by a specific population history of the Roma groups living in Croatia.

Until today, more than a dozen monogenic disorders have been identified in different groups of Roma. In most of them, specific founder mutations have been identified (Kalaydjieva *et al.*, 2001b, 2005; Kalaydjieva and Morar, 2003; Morar *et al.*, 2004). Disorders vary in the severity and age of onset as well as the organ systems they affect. Nevertheless, all of them have a serious impact on the health of an individual as well as

the whole group. In addition, they also represent a considerable financial burden for the health system and most often are not adequately recognized in that same system. It is important to note that some of these disorders are found in specific Roma groups, whereas others are found in several or all Roma groups. Certain conditions are also found in non-Roma populations, but mutations causing them are different than those found in the Roma, for example, GM1 gangliosidosis is also present in the Japanese population, but mutations causing it are different than the ones in the Roma (Nishimoto *et al.*, 1991; Yoshida *et al.*, 1991, 1992).

The Roma are a transnational population of common Indian origin living as a minority throughout most of the European countries. Their exodus from India occurred approximately a 1000 years ago, between the 5th and 10th centuries. After the journey through central Asia (Afghanistan and Persia), the Middle East, and present-day Turkey, they finally reached Europe around the 11th century. Upon arrival in Europe, a large part of the initial migrant population settled in the Balkan area, where their descendants still live today; another part continued their journey to western and northern Europe, and some of the migrants crossed the Danube and settled in Wallachia (present-day Romania), where they were enslaved. The slavery lasted for 500 years and the Roma that share this period of their history are today called the Vlach Roma. In the 19th century after the abolition of slavery, a new

TABLE 1. ROMA POPULATIONS INCLUDED IN THE STUDY AND NUMBER OF INDIVIDUALS TESTED FOR EACH DISEASE

Population/disease	Međimurje	Baranja	Total
HMSNL	162	265	427
GM1	161	261	422
CCFDN	150	100	250
LGMD2C	148	241	389

CCFDN, congenital cataracts, facial dysmorphism, and neuropathy; LGMD2C, limb girdle muscle dystrophy type 2C; HMSNL, hereditary motor and sensory neuropathy type Lom.

migration wave followed and these Vlax Roma, including the Bayash group among others, migrated to Croatia, Serbia, Hungary, and other Balkan states, as well as other parts of Europe and as far as the United States (Fraser, 1992; Hancock, 2002).

The period of slavery had a profound influence on the social structure of the Vlax Roma and during this time, the once homogenous group became separated into many different subgroups that started to differentiate among themselves (Fraser, 1992; Chaix *et al.*, 2004). Changes in the social structure combined with the practice of endogamy, eventually led to changes in the genetic structure, and as a result of this, present-day Vlax Roma groups have diverse genetic signatures. The migration routes they followed to escape slavery also add to the genetic differences between many groups, as do the early migrations they undertook to reach Europe. Recent studies of the mitochondrial DNA from various Roma groups show clear separation of the Vlax Roma from the Balkan and other Roma populations that reached Europe as part of the first migration wave (Mendizabal *et al.*, 2011; Peričić Salihović *et al.*, 2011). Similar results were shown by studies of autosomal (Gusmão *et al.*, 2008) and Y STR loci (Martinović Klarić *et al.*, 2009). Morar *et al.* (2004) indicated that the migration patterns are reflected in mutation distribution and different carrier rates of rare monogenic disorders in various Roma groups.

The goal of this study was to determine the genetic load of the Bayash Roma living in Croatia, and to explore how different migration routes influenced it, through the analysis of four single-gene disorders—hereditary motor and sensory neuropathy type Lom (HMSNL), GM1 gangliosidosis (GM1), congenital cataracts, facial dysmorphism, and neuropathy (CCFDN), and limb girdle muscle dystrophy type 2C (LGMD2C)—all caused by private mutations and noted in particular Roma groups, while absent in others.

Materials and Methods

The study included a total of 427 subjects who were tested for the presence of the four mutations, causing HMSNL, GM1, CCFDN, and LGMD2C diseases, to establish carrier frequencies. The exact number of subjects tested for the particular mutation is summarized in Table 1. The samples were collected as part of a multidisciplinary anthropological, molecular genetic and epidemiological research study of the Bayash Roma in Croatia (Škarić-Jurić *et al.*, 2007) at two locations, Međimurje in north-western Croatia and Baranja in eastern Croatia. These two populations differ between themselves based on the dialects they speak (Radosavljević, 2010). Informed consent has been obtained from all participants. The study protocol was approved by the Scientific Board and Ethical Committee of the Institute for Anthropological Research in Zagreb.

Diseases and founder mutations as well as detailed descriptions of the symptoms and molecular bases of the diseases have been provided elsewhere (Kalaydjieva *et al.*, 1996, 2000; Piccolo *et al.*, 1996; Angelicheva *et al.*, 1999; Tournev *et al.*, 1999; Merlini *et al.*, 2000; Morar *et al.*, 2004; Sinigerska *et al.*, 2006; Brunetti-Pierri and Scaglia, 2008).

Small samples of peripheral blood were collected for genetic analyses during fieldwork. DNA was extracted from white blood cells using a standard salting-out procedure (Miller *et al.*, 1988). PCR-based RFLP assays were used to detect the four mutations (Table 2) (Morar *et al.*, 2004; Sinigerska *et al.*, 2006). The digested products were resolved by agarose gel electrophoresis and visualized by SYBR safe staining. Each carrier was then confirmed by sequencing using a BigDye Terminator v3.1 Cycle Sequencing Kit and DNA Sequencing Analysis Software v3.7 on the ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Sequences were aligned and analyzed by the SeqScape v2.5 software (Applied Biosystems).

Allele frequencies were determined from the carrier rates.

Results

We identified a total of 4 individuals heterozygous for the mutation causing HMSNL, which was confirmed by direct sequencing. Carriers were detected in the Baranja population and were absent in the Međimurje population. The carrier rate in Baranja is 1.5%. Carriers for other three mutations causing GM1, CCFDN, and LGMD2C have not been found in our sample (Fig. 1).

Discussion

This study is the first systematic investigation of private disease-causing mutations in a sample of Roma people living

TABLE 2. PRIMERS AND RESTRICTION ENZYMES USED IN THE ANALYSES

Disease	Mutation site	Primers	Restriction enzyme
HMSNL	564C → T	HMSNL F: 5'-AAC TGT GGA GAA TAC GGG-3' HMSNL R: 5'-CTG TGC AGG CAG TTA CGG CAG C-3'	TaqI
GM1	176G → A	GM1 F: 5'-CAT CCT TGC TCT GTA GAA TG-3' GM1 R: 5'-ACA GTT GTA TCT TCT CTC CAG-3'	NlaIII
CCFDN	IVS6 + 389C → T	CCFDN F: 5'-CTT CCA GAG TTC ACG CCA TT-3' CCFDN R: 5'-ACA CAA AAG CCC AGC TCA AG-3'	NlaIII
LGMD2C	848G → A	LGMD2C F: 5'-CCT GTC TGT GGC CGG TGT GA-3' LGMD2C R: 5'-GCG TTT ACT TCC CAT CCA CGC TGC-3'	RsaI

For HMSNL, CCFDN, and LGMD2C according to Morar *et al.* (2004); for GM1 according to Sinigerska *et al.* (2006).

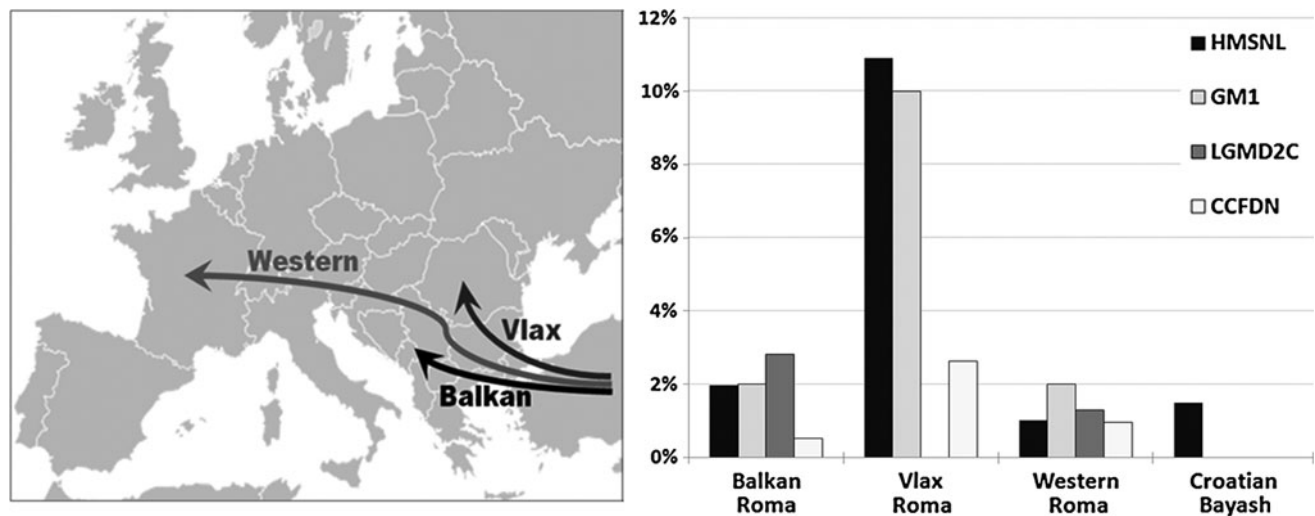


FIG. 1. Reported carrier rates for single-gene disorders in different migration categories and Croatian Bayash Roma with Roma migration paths illustrated on the left side. Carrier rates are given in percentages (according to Morar *et al.*, 2004). For GM1, carrier rates are for the general Roma population and Rudari subsolate in the Vlax migration category (Sinigerska *et al.*, 2006).

in Croatia, namely, the Bayash Roma subgroup, which is part of the bigger Vlax Roma group.

We identified a total of four individuals heterozygous for the mutation causing HMSNL in the population of Baranja (carrier rate is 1.5%). All known cases of HMSNL are of Romani origin, belonging to different Roma groups across Europe (Orpha.net, 2013); carrier rates vary from 0% to 16% in higher risk groups (Morar *et al.*, 2004). Morar *et al.* (2004) estimated the coalescent time in the entire Roma population for the mutation causing HMSNL to be ~850 years. This means that it predates the split of the Vlax Roma to numerous subgroups, which is dated to ~500 years (Kalaydjieva *et al.*, 2001a; Chaix *et al.*, 2004). The age and distribution of the mutation in most Roma groups implies its existence in the proto-Roma population, and we can assume that the mutation detected in the Croatian Bayash Roma founder population

represents a genetic legacy of this original proto-Roma group. Based on historical records, Hancock (1987) suggests that the enslavement of the Roma took place very soon after their arrival in Romania and forced separation resulted in lack of gene flow between different groups. This is supported by the work of Morar *et al.* (2004), who performed the haplotype analysis for the HMSNL mutation, which showed clustering by migration categories, but also revealed limited haplotype sharing between groups and the presence of unique haplotypes. In Vlax Roma groups, no haplotype was common to all of the three groups analyzed (Lom, Kalderash and Rudari) (Morar *et al.*, 2004). In the future, haplotype analysis of the detected carriers could also elucidate the origin of the mutation in Croatian Bayash and their relationship to other Vlax Roma groups, possibly even the exact group they descended from. Another possibility is that the mutation could have been

TABLE 3. INVESTIGATED DISORDERS, INCLUDING AFFECTED GENES, MUTATION SITES, REPORTED CARRIER RATES, AND AFFECTED ROMA GROUPS

Disease	Gene	Mutation site	Reported carrier rates	Roma groups where the disorder has been noted so far	Ref.
HMSNL	<i>NDRG1</i>	R148X	2% (16%) ^a	Balkan Roma (Kalaidjii North, Musicians, Xoroxane, Turgovzi)	Morar <i>et al.</i> (2004)
	8q24	564C → T		Vlax Roma (Lom ^a , Rudari, Kalderash)	
				Western Roma (Lithuanian, Spanish, Hungarian)	
GM1	<i>GLB1</i>	R59H	2% (10%) ^a	Vlax Roma (Rudari ^a)	Sinigerska <i>et al.</i> (2006)
	3p21.33	c.176G → A			
CCFDN	<i>CTDP1</i>	C → T substitution	1.4% (7%) ^a	Balkan Roma (Musicians, Feredjelli, Xoroxane),	Morar <i>et al.</i> (2004)
	18qter	389 bp downstream of the exon 6/ intron 6 junction		Vlax Roma (Kalderash, Rudari ^a)	
				Western Roma (Spanish)	
LGMD2C	γ -SG	C283Y	2% (6%) ^a	Balkan Roma (Turgovzi ^a , Darakchii ^a , Musicians)	Morar <i>et al.</i> (2004)
	13q12	848 G → A		Western Roma (Hungarian, Spanish)	

^aHigh-risk groups and rates.

introduced through intermarriage with Bayash Roma from Serbia (Chaix *et al.*, 2004), where it was also noted (Dačković *et al.*, 2008), and that the Croatian Bayash founder population originally did not have it. This could also explain the difference in carrier rates between the Međimurje and Baranja Bayash Roma populations. Nevertheless, more analysis is needed to corroborate either of these two hypotheses.

Another possibility is that the ancestors of the two Croatian Bayash groups became separated shortly after coming to Wallachia, where they were enslaved, implying that the Međimurje and Baranja groups were recently founded by two different Vlach groups, instead of one as previously presumed. Namely, it was already noted in the Y chromosome and mtDNA investigations that the Međimurje group is more homogenous, while the Baranja group is more diverse. They also comprised different Y and mtDNA haplotypes; moreover, the Baranja population has specific mtDNA haplogroups characteristic for India, which are not found in the Međimurje population (Martinović Klarić *et al.*, 2009; Peričić Salihović *et al.*, 2011), suggesting possible differentiation before arrival in Croatia. In addition to the genetic evidence, linguistic research has previously shown distinct dialectal differences between the two groups—the Baranja group speaks mostly the Muntean dialect of the Romanian language, while the Međimurje group uses Ardelean. Investigations have shown that these two dialects differ considerably between themselves, both on the phonological and morphological levels (Sorescu-Marinković, 2008; Radosavljević, 2010). In this light, differences in carrier rates could also be a result of this early separation of the ancestors of the two Croatian Bayash Vlach Roma groups, what was later enhanced by greater extent of genetic drift, maintained by a high degree of endogamy in the Međimurje population, compared with the Baranja group.

Other disease-causing mutations (LGMD2C, CCFDN, and GM1) investigated in this study have not been found in our sample, and the data on carrier rates are significantly different for various Roma groups (Morar *et al.*, 2004; Sinigerska *et al.*, 2006) (Table 3).

LGMD2C has so far been reported in Western Roma and different Balkan Roma groups living in Bulgaria, suggesting that the disorder occurs along the entire European migration route, including the Balkans and Central Europe (Kalaydjieva *et al.*, 2001b). Absence of the LGMD2C carriers in our sample is consistent with the results seen in other Vlach Roma groups, where the mutation was not noted in previous studies (Morar *et al.*, 2004).

CCFDN has been diagnosed in around 100 patients so far, who all belong to different migration categories of the Roma population and are homozygous for the same ancestral mutation (Kalaydjieva, 2006). Although Croatian Bayash are part of the Vlach Roma group, it is interesting that they do not carry the CCFDN mutation, whose carrier rate in the Vlach Roma is 2.64% while the age of the mutation is estimated to be between 400 to 650 years, which is consistent with the beginning of Roma slavery in Romania (Morar *et al.*, 2004). One of the possibilities for such a finding is that the group, where the CCFDN mutation occurred, was not in contact with the groups from which the Croatian Bayash eventually descended, or the mutation was originally present in Croatian groups, but became lost due to the effects of genetic drift.

GM1 mutation is present in low carrier frequencies in the general Roma population with higher frequencies in specific

Bulgarian Roma groups, mainly Rudari, which are highly endogamous (Sinigerska *et al.*, 2006). The mutation has not been detected in Croatian Bayash Roma.

However, when discussing the LGMD2C, CCFDN, and GM1 findings in Croatian Bayash Roma, we must not disregard the fact that Roma groups usually comprised a small number of individuals, which means that genetic forces such as drift can play a major role in shaping their genetic signatures. Therefore, based on our findings it is not yet possible to conclude a specific demographic history that would lead to today's genetic composition of the two Croatian groups, and more investigations are necessary for elucidating specific past episodes, which could influence the distribution of alleles specific for the monogenic disorders.

Our research has shown that there are distinct differences in carrier rates between two Croatian Bayash groups as well as compared with other Vlach groups; therefore, once again pointing to the significance of using the Roma group as a basic unit in the genetic research, as it was previously stated by Fraser (1992, 1998), where social and cultural features must not be disregarded in the interpretation of results obtained, as well as in the planning of the research itself.

Acknowledgments

We thank the numerous Gypsy individuals for their willingness to participate in this research. We are also especially grateful to Mr. Bajro Bajrić and his association "Roma for Roma" for generous help during field work. This research was supported by the Croatian Ministry of Science, Education and Sports, grant no. 196-1962766-2763 "Molecular-genetic portrait of the Roma—an isolated founder population model" to Branka Janičijević and 196-1962766-2747 "Complex trait variation and health in children, adults and centenarians" to Nina Smolej Narančić.

Author Disclosure Statement

No competing financial interests exist.

References

- Angelicheva D, Turnev I, Dye D, *et al.* (1999) Congenital cataracts facial dysmorphism neuropathy (CCFDN) syndrome: a novel developmental disorder in Gypsies maps to 18qter. *Eur J Hum Genet* 7:560–566.
- Brunetti-Pierri N, Scaglia F (2008) GM1 gangliosidosis. Review of clinical, molecular, and therapeutic aspects. *Mol Genet Metab* 94:391.
- Chaix R, Austerlitz F, Morar B, *et al.* (2004) Vlach Roma history: what do coalescent-based methods tell us? *Eur J Hum Genet* 12:285–292.
- Dačković J, Keckarević-Marković M, Komazec Z, *et al.* (2008) Hereditary motor and sensory neuropathy Lom type in a Serbian family. *Acta Myol* XXVII:59–62.
- Fraser A (1992) *The Gypsies*. Blackwell Publishers, Oxford.
- Fraser A (1998) *História do Povo Cigano*. Tradução de Costa T. Editorial Teorema, LDA.
- Gusmão A, Gusmão L, Gomes V, *et al.* (2008) A perspective on the history of the Iberian gypsies provided by phylogeographic analysis of Y-chromosome lineages. *Ann Hum Genet* 72:215–227.
- Hancock I (1987) *The Pariah Syndrome: An Account of Gypsy Slavery and Persecution*. Karoma Publishers, Ann Arbor, MI.

- Hancock I (2002) We Are Romani people. Ames and e Romane džene. University of Hertfordshire Press, Hatfield.
- Kalaydjieva L, Calafell F, Jobling MA, *et al.* (2001a) Patterns of inter- and intragroup genetic diversity in the Vlax Roma as revealed by Y chromosome and mitochondrial DNA lineages. *Eur J Hum Genet* 9:97–104.
- Kalaydjieva L, Gresham D, Calafell F (2001b) Genetic studies of the Roma (Gypsies): a review. *BMC Med Genet* 2:5.
- Kalaydjieva L, Gresham D, Gooding R, *et al.* (2000) N-myc downstream-regulated gene 1 is mutated in hereditary motor and sensory neuropathy–Lom. *Am J Hum Genet* 67:47–58.
- Kalaydjieva L, Hallmayer J, Chandler D, *et al.* (1996) Gene mapping in Gypsies identifies a novel demyelinating neuropathy on chromosome 8q24. *Nat Genet* 14:214–217.
- Kalaydjieva L (2006) Congenital cataracts–facial dysmorphism–neuropathy. *Orphanet J Rare Dis* 1:32.
- Kalaydjieva L, Morar B (2003) Roma (Gypsies): genetic studies. In: Cooper DN (eds) *Nature Encyclopedia of the Human Genome*, Volume 5. Nature Publishing Group, London, pp 160–165.
- Kalaydjieva L, Morar B, Chaix R, *et al.* (2005) A newly discovered founder population: the Roma/Gypsies. *BioEssays* 27: 1084–1094.
- Martinović Klarić I, Peričić Salihović M, Barać Lauc L, *et al.* (2009) Dissecting the molecular architecture and origin of Bayash Romani patrilineages: genetic influence from South Asia and the Balkans. *Am J Phys Anthropol* 138:333–342.
- Mendizabal I, Calente C, Gusmao A, *et al.* (2011) Reconstructing the Indian origin and dispersal of the European Roma: a maternal genetic perspective. *PLoS One* 6:e15988.
- Merlini L, Kaplan JC, Navarro C, *et al.* (2000) Homogeneous phenotype of the Gypsy limb-girdle MD with the gamma-sarcoglycan C283Y mutation. *Neurology* 54:1075–1079.
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215.
- Morar B, Gresham D, Angelicheva D, *et al.* (2004) Mutation history of the Roma/Gypsies. *Am J Hum Genet* 75:596–609.
- Nishimoto J, Nanba E, Inui K, *et al.* (1991) GM1-gangliosidosis (genetic beta-galactosidase deficiency): identification of four mutations in different clinical phenotypes among Japanese patients. *Am J Hum Genet* 49:566–574.
- Orpha.net (2013) The portal for rare diseases and orphan drugs. Available at www.orpha.net/consor/cgi-bin/index.php?lng=EN (accessed June 11 2013).
- Peričić Salihović M, Barešić A, Martinović Klarić I, *et al.* (2011) The role of the Vlax Roma in shaping the European Romani maternal genetic history. *Am J Phys Anthropol* 146:262–270.
- Piccolo F, Jeanpierre M, Leturcq F, *et al.* (1996) A founder mutation in the gamma-sarcoglycan gene of gypsies possibly predating their migration out of India. *Hum Mol Genet* 5: 2019–2022.
- Radosavljević P (2010) Jezik Roma Bajaša na teritoriju Republike Hrvatske. PhD thesis. University of Zagreb, Faculty of Humanities and Social Sciences.
- Sinigerska I, Chandler D, Vaghjani V, *et al.* (2006) Founder mutation causing infantile GM1-gangliosidosis in the Gypsy population. *Mol Genet Metab* 88:93–95.
- Škarić-Jurić T, Martinović Klarić I, Smolej Narančić N, *et al.* (2007) Trapped between tradition and transition—anthropological and epidemiological cross-sectional study of Bayash Roma in Croatia. *Croat Med J* 48:708–719.
- Sorescu-Marinković A (2008) The Bayash in Croatia: Romanian vernaculars in Baranja and Medjimurje. In: Sikimić B, Ašić T (eds) *The Romance Balkans*. The Institute for Balkan Studies, Belgrade, Serbia, pp. 173–225.
- Tournev I, Kalaydjieva L, Youl B, *et al.* (1999) Congenital cataracts facial dysmorphism neuropathy syndrome, a novel complex genetic disease in Balkan Gypsies: clinical and electrophysiological observations. *Ann Neurol* 45:742–750.
- Yoshida K, Oshima A, Sakuraba H, *et al.* (1992) GM1 Gangliosidosis in adults: clinical and molecular analysis of 16 Japanese patients. *Ann Neurol* 31:328–332.
- Yoshida K, Oshima A, Shimamoto M, *et al.* (1991) Human beta-galactosidase gene mutations in G(M1)-gangliosidosis: a common mutation among Japanese adult/chronic cases. *Am J Hum Genet* 49:435–442.

Address correspondence to:
 Marijana Peričić Salihović, PhD
 Institute for Anthropological Research
 Ulica Ljudevita Gaja 32
 10000 Zagreb
 Croatia (Hrvatska)
 E-mail: mperic@inantro.hr