MOLECULAR AND EXPERIMENTAL BIOLOGY IN MEDICINE

ORIGINAL ARTICLE

THE VARIABILITY OF MULTI-DRUG RESISTANCE *ABCB1* GENE IN THE ROMA POPULATION FROM CROATIA

Matea Zajc Petranovic, Zeljka Tomas, Tatjana Skaric-Juric, Nina Smolej Narancic, Branka Janicijevic, Anita Stojanovic Markovic, Marijana Pericic Salihovic

Abstract: When overexpressed, a large transmembrane P-glycoprotein, the product of the ABCB1 gene, is a notable impediment to brain-targeted therapies (like antiepileptics) and chemotherapies. Some of the genetic biomarkers with evidence of multi-drug resistance in ABCB1 — rs1045642, rs1128503, and rs3213619 — were analyzed in 440 subjects, members of three socio-culturally different Roma (Gypsy) groups of Croatia.
Minor allele frequencies (MAFs) of rs1045642 and rs1128503 were the highest in the Balkan Roma (63.6% and 69.4%, respectively) when compared to the Baranja (52.3% and 62.5%) and the Međimurje Roma (48.8% and 54.5%) ($p=0.0005$ and $p=0.0011$, respectively). rs3213619 was monomorphic in the Međimurje group, while its MAFs in other two Roma groups were very low (<1.9%). The distribution of five detected haplotypes (four in the Međimurje group) significantly differed between the Roma subpopulations ($p<0.0001$), just like the frequencies of diplotypes ($p=0.0008$). At a global scale, the positive relationship between genetic and geographic distances between the 21 investigated populations indicates isolation by spatial distance. However, this is not true for the relationship between Roma and other populations due to their population history.
The analyzed ABCB1 loci indicate genetic distinctiveness of the Roma population.

Institute for Anthropological Research, Zagreb, Croatia

Corresponding author:
Matea Zajc Petranovic
Institute for Anthropological Research
Ljudevita Gaja 32,Zagreb, Croatia
Tel: +385 1 55 35 136; Fax: +385 1 55 35 105
e-mail: matea@inantro.hr

Submitted: December, 2018
Accepted: March, 2019

Key words: *ABCB1* gene, rs1045642, rs1128503, rs3213619, pharmacogenetics, Roma (Gypsy), Croatia

INTRODUCTION

Studies of European genetic diversity continue to untangle how people settled and migrated through Europe and adapted to their environment.[1] Although most of the present-day European populations are well-studied, some isolated populations, like the Roma (Romani, Gypsy), are not that well known yet.
The Roma, a large transnational minority population numbering more than 15 million people worldwide, 10-12 million of whom live in Europe, originate from India. The first Roma left India sometime between the 5th and the 10th century and reached Asia Minor (present day Turkey) in the 11th century. A large part of the initial migrant population settled in South-Eastern Europe (the Balkans) where their descendants, the Balkan Roma, still live today, while some Roma continued their journey to Western and Northern Europe around 1400s. A part of the latter Roma crossed the Danube river and ended up in Wallachia (present day Romania), where they were enslaved and forced to work in mines until 1850, when slavery was abolished. After the abolition of slavery these Vlax Roma (also called Bayash Roma because they speak ljimba d`byash) migrated to Serbia, Hungary, Croatia and other countries. Both Balkan and Bayash Roma are examples of population isolates with persistent, centuries-long socio-cultural and reproductive isolation,[2, 3] which is reflected in their genetic structure.[4-8]
The adenosine triphosphate (ATP)-binding cassette sub-family B member 1 gene (ABCB1), also known as multi-drug resistance protein 1 (MDRP1) gene, is one of many ubiquitous ATP-binding cassette (ABC) genes. ABC transporters are a superfamily of large

membrane proteins which are able to transport a variety of compounds through membranes against steep concentration gradients at the cost of ATP hydrolysis.[9] Many ABC genes were originally discovered during the positional cloning of human genetic disease genes - 14 ABC genes were found to be linked to Mendelian disorders, some of them with complex diseases as well.[10, 11]

The *ABCB1* gene, located on the minus strand of chromosome 7q21.12, is 209,691 bases long, consists of 29 exons and is highly polymorphic (with more than 100 polymorphic sites with minor allele frequency >5%), although the majority of its SNPs are either intronic or non-coding.[12] *ABCB1* encodes human P-glycoprotein (P-gp) (1280 amino acids), a large transmembrane protein expressed in a polarized manner in the plasma membrane of cells in barrier and elimination organs, where it has protective and excretory functions.[13] In different organs P-gp has different functions: in the small intestine, colon, and bile-facing canaliculi of liver it eliminates orally administered drugs thus limiting their bioavailability, while from the kidney, and again through biliary excretion, it eliminates substrates. Furthermore, P-gp restricts the permeability of drugs from blood into the brain, cerebral spinal fluid, placenta and testes.[14] In immunological and other blood components, P-gp plays a role in viral resistance and in trafficking cytokines and enveloped viruses.[15, 16]

Many studies have investigated the role of P-gp compounds in altering the pharmacokinetics of administered drugs, resulting in harmful as well as beneficial effects: some interactions affect drug safety and efficacy,[17] while others optimize drug delivery. Multi-drug resistance from intrinsic (drug-naive) and acquired (drug-induced) overexpression of P-gp is a strong barrier to brain-targeted drugs (like antiepileptics and neuroantiretrovirals) and chemotherapies (like in acute myeloid leukemia, sarcoma and other cancers).[12] On the Core Marker List of evidence-based drug metabolizing (ADME) genetic biomarkers,[18] there are four loci in the *ABCB1* gene: rs3213619 (-129T>C), rs1128503 (1236T>C), rs2032582 (2677T>G/A) and rs1045642 (3435C>T). The three biallelic loci analyzed in this study will be presented in more detail.

The rs3213619 (-129T>C) SNP located in exon 1b (promotor region) of P-gp, showed significant association with taxane-related sensory neuropathy after chemotherapy treatments of patients with early breast cancer.[19] The rs1128503 (1236T>C), a synonymous SNP, encodes for the TM6 region of P-gp, which is essential for substrate binding. This polymorphism is associated with tumor response to chemotherapy in cancer patients[20] and implicated in lovastatin and clopidogrel response modulation when treating hypercholesterolemia and reducing the risk of heart disease and stroke.[21] The rs1045642 (3435C>T) modifies the gene expression of P-gp without altering the sequence of the protein (wobble mutation), probably by altering mRNA stability[22] as well as the folding of the protein. The *ABCB1* 3435C>T polymorphism is associated with reduced bioavailability and increased risk of cardiovascular death for patients taking clopidogrel.[23] The latter two coding *ABCB1* SNPs, in combination with 2677T>G/A not analyzed here, are in high linkage disequilibrium.[24]

The aims of this study were: (1) to determine the allele, genotype and haplotype frequencies of three clinically relevant *ABCB1* SNPs in three Roma groups living in Croatia, (2) to compare ABCB1 alleles of the Croatian Roma with the majority Croatian population, European and global populations, and (3) to evaluate the position of the Croatian Roma in the global genetic landscape.

MATERIAL AND METHODS

DNA samples of 440 adult individuals were collected during the on-going multidisciplinary anthropological, molecular-genetic and epidemiological investigation of Roma populations in Croatia. Study participants belonged to three socio-culturally different Roma groups: two groups of Vlax (Bayash) Roma, the Baranja group and the Međimurje group, and one Balkan Roma group. The Roma participated in the study voluntarily and were informed about the goals, methods and expectations of the study with the help of linguistically and culturally competent and trained Roma volunteers. The research protocol was approved by the Ethical Committee of the Institute for Anthropological Research in Zagreb, Croatia.

The genotyping of three single nucleotide polymorphisms (SNPs) in the *ABCB1* gene (rs1045642, rs1128503 and rs3213619) was carried out using the KASP method.[25] These three SNPs were chosen for the analysis since they are on the list of evidence-proved biomarkers associated with the metabolism of drugs (ADME Core Marker List).[18] Allele and genotype frequencies were calculated by the direct counting method. Hardy-Weinberg equilibrium (HWE) and exact test of population differentiation were performed using Arlequin 3.5.2.2. The genotype and allele frequency differences between Roma groups were tested using the Chi-square test. The analyses were performed using the SPSS Statistics 11.0 statistical package for Windows (SPSS Inc., Chicago, IL, USA), with statistical significance set at $P<0.05$. Haplotypes were inferred using Phase ver. 2.1 but could not be named according to *ABCB1* star allele designations since they are not harmonized in the literature.[12]

The genetic distance matrix between the Croatian Roma population and 20 populations with different genetic ancestry from the 1000 Genomes Project Phase 3 list was computed according to the method of Nei (1972) using Poptree2 software.[26] These 20 populations belong to five large continental regions, and each region was represented by four populations: (1) European region (EUR): Finland, Italy, Spain, UK, (2) South Asian region (SAS): Bangladesh, India,

Pakistan, Sri Lanka, (3) African region (AFR): Gambia, Kenya, Nigeria, Sierra Leone, (4) Central and South American region (AMR): Colombia, Mexico, Peru, Puerto Rico, and (5) Eastern Asian region (EAS): Dai Chinese, Han Chinese, Japan, Vietnam (China was represented by two distinct populations, Han and Dai; Han is a majority population while Dai is here representing non-Han China populations). In addition to 1000 Genomes data, we enlarged the sample by including the genotyping data from publications citing any of the three investigated SNPs in the above mentioned 20 populations (described in detail in a paper by Škarić-Jurić et al[27]). The selection criteria for the usage of data from these publications were: (1) a clearly stated study population and, where relevant, participants' ethnicity, (2) allele frequency and (3) the number of participants in the study. These additional data enlarged the size of seven 1000 Genomes populations: Italy, Spain, UK, India, Kenya, Japan and Han Chinese. Allele frequencies were calculated by weighting samples for each population.

Statistical comparison of allele frequencies (proportions) was performed in MedCalc online statistical software.[28]

The Mantel test of correlation between genetic and geographic distances was performed using non-commercial software IBD: *Isolation by distance* v1.52 (available at http://www.bio.sdsu.edu/pub/andy/IBD.html). Geographic distances between the analyzed populations were calculated using two types of free online software: *iTouchMap* and *Movable Type Scripts*. *iTouchMap* calculates latitude and longitude of a point, and *Movable Type Scripts* calculated distance between latitude/longitude points (available at http://itouchmap.com/latlong.html and http://www.movable -type.co.uk/scripts/latlong.html).

RESULTS

In comparison with the weighted averages of the 1000 Genomes global 20 populations, the MAFs of rs1045642, rs1128503 and rs3213619 in the Roma significantly differed (Table 1). The MAF of rs3213619 was significantly lower in the Croatian Roma than in the 1000 Genomes populations` average (p=0.0004), while the MAFs of rs1045642 and rs1128503 were higher in the Roma than in the 1000 Genomes global populations (both p<0.0001). When compared to the 1000 Genomes European subset, MAF of rs3213619 was again significantly lower (p=0.0056), while rs1128503 MAF was significantly higher (p<0.0001) in the Roma. The rs1045642 MAF did not differ between the Croatian Roma and the 1000 Genomes European populations. In comparison with the surrounding majority Croatian population,[29] the MAF of rs1045642 in the Roma was similar to that in Croats (55.9% vs. 55.1%, p=ns), while MAF of rs1128503 was significantly more frequent in the Roma sample than in Croats (62.9% vs. 46.0%, p=0.0015). The rs3213619 was not investigated in that Croatian study.

The results of genotyping of three investigated SNP loci in the DNA samples of the three Croatian Roma groups are given in Table 2. Loci rs1128503 and rs1045642 in the Međimurje and in the Balkan Roma were in Hardy-Weinberg equilibrium, as opposed to the Baranja Roma. An exact test of population differentiation showed that the three Roma subpopulations were differentiated when analyzing loci rs1045642 and rs1128503 (p=0.00000±0.00000, 30000 Markov steps done). Locus rs3213619 was not included in the HWE calculations since its MAF was <0.05 in all three Roma subpopulations. *ABCB1* SNPs rs1045642 genotype and allele frequencies significantly differed between the Baranja and the Balkan Roma, and between the Međimurje and the Balkan Roma. Locus rs1128503 genotype and allele frequencies significantly differed only between the Međimurje and the Balkan Roma. Both of these *ABCB1* polymorphic loci showed significant differences between all three Roma subpopulations in genotype as well as in allele frequencies distribution, with A alleles in both SNPs being the most frequent in

Table 1. Single nucleotide polymorphisms (SNPs) of the ABCB1 gene analyzed in this study. Loci are sorted according to the ascending chromosome position on the forward strand, but the gene is negative stranded. Only successfully genotyped SNPs were counted in column N. The allele defined as minor in the Croatian Roma refers to the global minor allele in 1000 Genomes' database. The statistically significant differences between the Croatian Roma and other populations were determined using difference in proportions test.[28]

SNP	1000 Genomes major allele	1000 Genomes minor allele	1000 Genomes minor allele frequency ALL	1000 Genomes minor allele frequency Europe	Croatia minor allele frequency (reference 29)	Croatian Roma minor allele frequency (this study)	Statistical comparison of allele frequencies
rs1045642	G	A	40% (N=5008)	52% (N=1006)	55.1% (N=99)	55.9% (N=435)	a
rs1128503	G	A	42% (N=5008)	42% (N=1006)	46.0% (N=107)	62.9% (N=421)	b, c, d
rs3213619	A	G	5% (N=5008)	4% (N=1006)	-	1.2% (N=426)	e, f

a Croatian Roma vs. 1000 Genomes ALL populations; $\chi2=41.913$, df=1, p<0.0001
b Croatian Roma vs. 1000 Genomes ALL populations; $\chi2=69.264$, df=1, p<0.0001
c Croatian Roma vs. 1000 Genomes European populations; $\chi2=52.063$, df=1, p<0.0001
d Croatian Roma vs. Croatian population; $\chi2=10.102$, df=1, p=0.0015
e Croatian Roma vs. 1000 Genomes ALL populations; $\chi2=12.707$, df=1, p=0.0004
f Croatian Roma vs. 1000 Genomes European populations; $\chi2=7.677$, df=1, p=0.0056

Table 2. ABCB1 genotype and allele frequencies (%) in the Croatian Roma for three subpopulations (Baranja, Međimurje and Balkan). Genotyped loci are listed according to their genomic positions in Chromosome 7. The differences in frequencies of genotypes and alleles in polymorphic loci between the three subpopulations were tested using the Chi-square test.

SNP	genotype / allele	Baranja		Međimurje		Balkan		Baranja vs. Međimurje	Baranja vs. Balkan	Međimurje vs. Balkan	Baranja vs. Međimurje vs. Balkan
		N	%	N	%	N	%				
rs1045642	GG	131	19.1	128	24.2	176	14.8	ns	χ^2=11.282, df=2, p=0.0036	χ^2=14.299, df=2, p=0.0008	χ^2=19.640, df=4, p=0.0006
	AG		57.3		53.9		43.2				
	AA		23.7		21.9		42.0				
	A	137	52.3	125	48.8	224	63.6	ns	χ^2=7.982, df=1, p=0.0047	χ^2=13.291, df=1, p=0.0003	χ^2=15.121, df=2, p=0.0005
rs1128503	GG	128	10.2	123	19.5	170	7.6	ns	ns	χ^2=14.466, df=2, p=0.0007	χ^2=16.604, df=4, p=0.0023
	GA		54.7		52.0		45.9				
	AA		35.1		28.5		46.5				
	A	160	62.5	134	54.5	236	69.4	ns	ns	χ^2=13.689, df=1, p=0.0002	χ^2=13.690, df=2, p=0.0011
rs3213619	GG	130	0	126	0	170	0	ns	ns	ns	ns
	GA		3.8		0		2.9				
	AA		96.1		100		97.1				
	G	5	1.9	0	0.0	5	1.5	—	—	—	—

N - number of individuals; ns - not significant

Table 3. The frequency (%) and number (N) of ABCB1 haplotypes in the total Croatian Roma population and separately for the three Roma subpopulations (Baranja, Međimurje, and Balkan). The differences in haplotype frequencies between subpopulations were tested using the Chi-square test.

No.	Haplotypes*	Total N (%)	Baranja N (%)	Međimurje N (%)	Balkan N (%)	Baranja vs. Međimurje	Baranja vs. Balkan	Međimurje vs. Balkan	Baranja vs. Međimurje vs. Balkan
1	AAA	476 (54.2)	135 (51.1)	126 (48.1)	215 (60.7)	$\chi2$ = 14.434, df = 4, p = 0.0060	ns	$\chi2$ = 27.775, df = 4, p < 0.0001	$\chi2$ = 34.298, df = 8, p < 0.0001
2	GGA	308 (35.0)	92 (34.9)	120 (45.8)	96 (27.1)				
3	GAA	71 (8.1)	29 (11.0)	14 (5.3)	28 (7.9)				
4	GGG	10 (1.1)	5 (1.9)	0	5 (2.8)				
5	AGA	15 (1.7)	3 (1.1)	2 (0.8)	10 (1.4)				
Total		880 (100.0)	264 (100.0)	262 (100.0)	354 (99.9)				

*loci order: rs1045642, rs1128503, and rs3213619

the Balkan Roma (Table 2).

SNP genotyping enabled the defining of five distinct haplotypes for the loci ordered according to the ascending chromosome position on the forward strand: rs1045642, rs1128503, and rs3213619 (Table 3). All five haplotypes were present in the Baranja and in the Balkan Roma, although with different frequencies, while in the Međimurje Roma only four haplotypes

Table 4. Diplotype combinations detected in the Croatian Roma. The differences in diplotype frequencies between subpopulations were tested using the Chi-square test but only for diplotypes with the number of subjects per cell ≥5 (diplotypes 1,2,4,7 and 9).

No.	Diplotypes	Total N (%)	Baranja N (%)	Međimurje N (%)	Balkan N (%)	Baranja vs. Međimurje	Baranja vs. Balkan	Međimurje vs. Balkan	Baranja vs. Međimurje vs. Balkan
1	GGA/GGA	49 (11.1)	12 (9.1)	25 (19.1)	12 (6.8)	ns	ns	$\chi2$ = 21.690, df = 4, p = 0.0002	$\chi2$ = 26.800, df = 8, p = 0.0008
2	GGA/GAA	22 (5.1)	6 (4.6)	5 (3.8)	11 (6.2)				
3	GGA/AGA	5 (1.1)	2 (1.5)	1 (0.8)	2 (1.1)				
4	GGA/AAA	183 (41.6)	60 (45.5)	64 (48.9)	59 (33.3)				
5	GAA/GGG	5 (1.1)	3 (2.3)	0	2 (1.1)				
6	GAA/GAA	6 (1.4)	4 (3.0)	1 (0.8)	1 (0.6)				
7	GAA/AAA	32 (7.3)	12 (9.1)	7 (5.3)	13 (7.3)				
8	AGA/AAA	10 (2.3)	1 (0.8)	1 (0.8)	8 (4.5)				
9	AAA/AAA	123 (27.9)	30 (22.7)	27 (20.6)	66 (37.3)				
10	AAA/GGG	5 (1.1)	2 (1.5)	0	3 (1.7)				
Total		440 (100.0)	132 (100.1)	131 (100.1)	177 (99.9)				

*loci order: rs1045642, rs1128503, and rs3213619

Table 5. The pairwise differences in Nei`s genetic distances adjusted for sample size (DST) values between 21 populations.

Populations																							
Populations		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
	1	0	0.040	0.023	0.063	0.050	0.015	0.020	0.034	0.025	0.208	0.277	0.251	0.280	0.065	0.029	0.071	0.068	0.047	0.025	0.054	0.038
	2		0	0.027	0.044	0.038	0.044	0.049	0.045	0.052	0.146	0.200	0.181	0.207	0.055	0.028	0.043	0.054	0.080	0.056	0.078	0.060
	3			0	0.036	0.032	0.031	0.032	0.031	0.038	0.134	0.189	0.166	0.195	0.043	0.016	0.032	0.043	0.053	0.030	0.052	0.036
	4				0	0.046	0.072	0.069	0.056	0.078	0.101	0.146	0.125	0.153	0.052	0.032	0.029	0.049	0.083	0.061	0.077	0.060
	5					0	0.056	0.058	0.051	0.063	0.131	0.182	0.162	0.188	0.056	0.032	0.040	0.054	0.081	0.058	0.078	0.061
	6						0	0.025	0.041	0.027	0.233	0.304	0.280	0.308	0.075	0.038	0.085	0.079	0.059	0.036	0.067	0.051
	7							0	0.042	0.035	0.209	0.275	0.250	0.278	0.071	0.037	0.076	0.074	0.054	0.032	0.060	0.045
	8								0	0.049	0.163	0.220	0.198	0.224	0.061	0.033	0.056	0.062	0.063	0.041	0.064	0.049
	9									0	0.231	0.300	0.276	0.303	0.081	0.045	0.089	0.084	0.066	0.044	0.073	0.057
	10										0	0.073	0.052	0.080	0.116	0.116	0.057	0.104	0.185	0.167	0.163	0.148
	11											0	0.071	0.096	0.161	0.169	0.097	0.148	0.243	0.227	0.217	0.203
	12												0	0.078	0.141	0.146	0.075	0.127	0.216	0.199	0.190	0.177
	13													0	0.167	0.175	0.105	0.154	0.244	0.229	0.218	0.205
	14														0	0.040	0.041	0.058	0.082	0.062	0.077	0.062
	15															0	0.025	0.039	0.053	0.031	0.051	0.035
	16																0	0.035	0.080	0.059	0.070	0.054
	17																	0	0.084	0.063	0.078	0.062
	18																		0	0.041	0.062	0.050
	19																			0	0.044	0.031
	20																				0	0.051
	21																					0

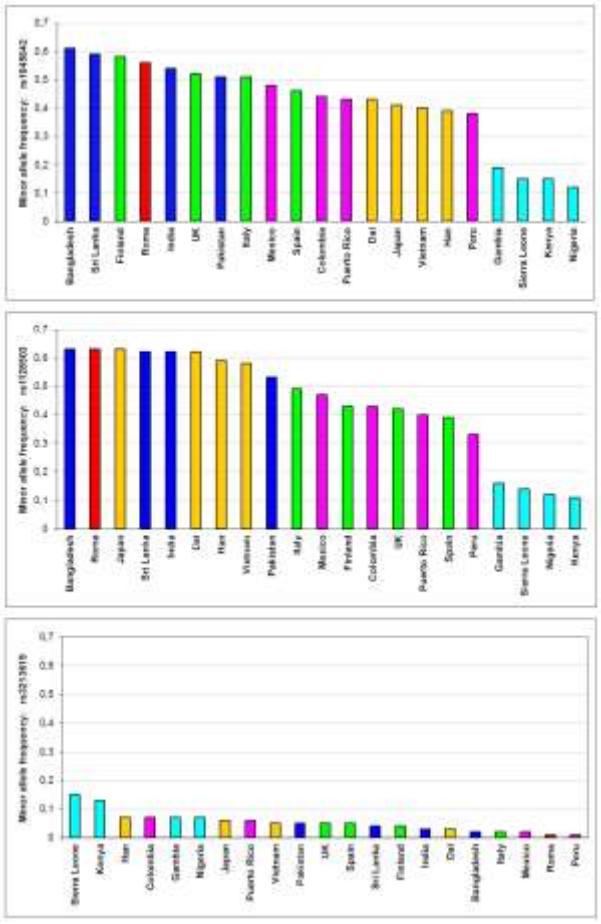


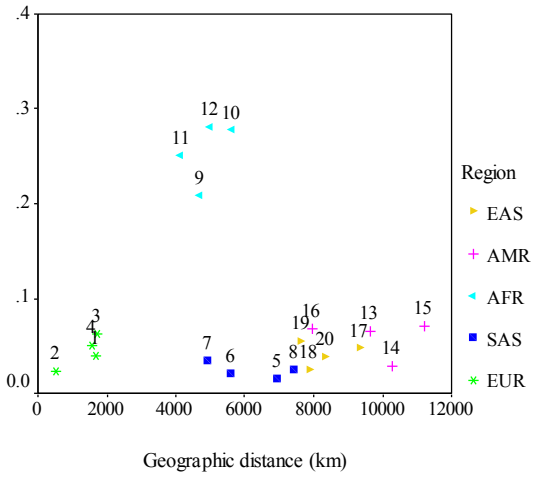
Figure 1. The global minor allele frequencies for the three investigated ABCB1 SNP loci in 21 populations sorted in descending order. Continental regions are marked as follows: European - green, South Asian - blue, Eastern Asian - yellow, Central and South American - magenta, and African - turquoise. Croatian Roma population is marked in red.

were detected. The haplotype AAA was the most frequent, and haplotype AGA was the least frequent in all three populations. Haplotype GGG, which was the second least frequent in the Baranja and in the Balkan Roma, was not detected in the Roma from Međimurje. The *ABCB1* haplotype frequencies significantly differed between the Baranja and the Međimurje Roma ($p=0.0060$) and between the Međimurje and the Balkan Roma ($p<0.0001$). Chi-square test results also showed significant differences in haplotype frequencies between all the three Roma subpopulations ($\chi 2 = 34.298$; $df = 8$; $p < 0.0001$).

ABCB1 gene diplotype combinations in the Croatian Roma are presented in Table 4. A total of ten diplotype combinations were detected in the Baranja and in the Balkan Roma, while eight diplotypes were detected in the Međimurje Roma. The diplotype GGA/AAA, which was the most frequent in the Baranja (45.45%) and in the Međimurje Roma (48.86%), was the second most frequent in the Balkan Roma (33.33%). On the other hand, the diplotype AAA/AAA, which was the most frequent in the Balkan Roma (37.29%), was the second most frequent in the Baranja and in the Međimurje sample (22.73% and 20.61%, respectively). The significant differences in diplotype frequencies between subpopulations were confirmed using the Chi-square test between the Balkan and the Međimurje Roma ($p=0.0002$) and between all three investigated Croatian Roma populations ($p=0.0008$). Only diplotypes with number of subjects per cell ≥5 were included in testing (diplotypes 1, 2, 4, 7, and 9).

The degree of similarity in the genetic structure between the Croatian Roma and other investigated 1000 Genomes populations was evaluated from Nei`s genetic distances corrected for sample size, which ranged from 0.015 (between the Croatian Roma and Bangladesh populations) to 0.308 (between Bangladesh and Sierra Leone populations) (Table 5). Comparing Croatian Roma to other twenty populations, besides the already mentioned lowest level of their differentiation from the Bangladesh population, the greatest divergence was observed when the Roma were compared with populations from the African region: 0.280 with the Sierra Leone population, 0.277 with the population of Kenya, 0.251 with the population Nigeria, and 0.208 with the population of Gambia.

The global minor allele frequencies for 3 SNPs were presented graphically (Figure 1). The African region populations had the most extreme MAF values when compared to the Roma and the other 1000 Genomes populations; in rs1045642 and rs1128503, all African populations had the lowest MAF values, while in rs3213619 two populations` (Kenya and Sierra Leone) had the highest MAFs, higher than all other analyzed population.



Legend: 1 = Finland; 2 = Italy; 3 = Spain; 4 = UK; 5 = Bangladesh; 6 = India; 7 = Pakistan; 8 = Sri Lanka; 9 = Gambia; 10 = Kenya; 11 = Nigeria; 12 = Sierra Leone; 13 = Colombia; 14 = Mexico; 15 = Peru; 16 = Puerto Rico; 17 = Japan; 18 = China_Dai; 19 = China_Han; 20 = Vietnam.

Figure 2. Genetic distances between the Croatian Roma and 20 populations worldwide in relation to their geographic distances. These 20 populations are grouped according to five large continental regions: European region (EUR), South Asian region (SAS), African region (AFR), Central and South American region (AMR), and Eastern Asian region (EAS).

The relationship between genetic variation and geographic distance of twenty-one investigated populations was analyzed using the correlation between genetic and geographic distances` matrices. The correlation was positive and significant (Pearson's r=0.391, p=0.001 after 1,000 permutations) indicating isolation by spatial distance at the global scale. Still, pairwise comparisons of genetic and geographical distances between the Croatian Roma and twenty 1000 Genomes populations showed no correlation (Pearson's r= -0.058, p=0.505) (Figure 2), not even in case when the genetically most distant African populations were excluded from the correlation analysis (Pearson's r= -0.032, p=0.477).

DISCUSSION

The genetic distinctiveness of the Roma population when considering pharmacogenes has been detected in various individual genes; in some of the *CYP* genes,[30-32] in *SLCO1B1,*[33] *VKORC1,*[34] *PON1* and *P2RY12*[35] genes. In addition, it was recently confirmed in a comprehensive analysis of a set of 95 pharmacogenomically relevant polymorphic markers (from 31 core *ADME* genes) of drug response and adverse drug reactions in the Roma and twenty other, ethnically different populations from the 1000 Genomes database.[27]

The comparison of *ABCB1* genotypes and alleles of three polymorphic loci between the three Croatian Roma groups showed that the Vlax Roma (from Baranja and Međimurje) significantly differ from the non-Vlax, Balkan Roma. Furthermore, the Roma from Međimurje, in comparison with the other two Roma groups, had the lowest genetic diversity. Due to the absence of the G minor allele in rs3213619, the Roma from Međimurje also do not have one *ABCB1* haplotype (GGG) and two *ABCB1* diplotypes (GAA/GGG and AAA/GGG) which were found in the other two Croatian Roma groups. These results go in line with the mitochondrial DNA analysis results, which revealed that the distribution of haplogroups in the Međimurje Roma was less diverse than in the Baranja Roma. The diversity indexes also pointed to a lower level of variation among the Vlax Roma settled in Međimurje than in Baranja.[36]

The Roma's significant difference in allele frequencies from the surrounding, majority Croatian population, also confirms their genetic specificity which is most probably the consequence of their practice of endogamy. It is particularly true for rs1128503 and rs3213619 loci. On the other hand, high MAF of the rs1045642 locus only characterizes the Balkan Roma group, while the MAFs of both Croatian Bayash groups compare well with MAFs in other global populations, including the European 1000 Genomes subset, and the Croatian population.

The *ABCB1* functional haplotypes consist of at least three coding loci from the ADME Core list.[37-40] Unfortunately, since the triallelic rs2032582 (2677T>G/A) could not be genotyped using the KASP method, the haplotype frequencies or the functional implications of the *ABCB1* haplotypes could not be compared to literature data on other populations.

Genetic distances of the three investigated ABCB1 variants between the Croatian Roma and other twenty populations with different ethnic origin positioned Roma equally close to the South Asian population (SAS) and to the European populations (EUR), despite being geographically several times further to SASs (5000-7500 km) than to EURs (500-1500 km), while the African region populations were the furthest. The genetic closeness with the South Asian populations was also found in several previous studies of various genetic markers including 95 *ADME* loci,[27] mitochondrial DNA and Y chromosome markers[41, 42] and the genome-wide data,[43, 44] proving the Indian ancestry of the Roma.

The clustering of the *ABCB1* polymorphisms investigated here differs from the results of the comprehensive study of 95 *ADME* loci[27] mentioned earlier, in close relation of the former to East Asian and American regions populations, although spatially even further than the SAS populations. Considering that the 95 *ADME* loci analysis[27] included many pharmacogene biomarkers and was conducted on a much larger sample, its results are more trustworthy in gaining the complete picture on the position of the Roma within the global ADME genetic landscape.

Still, the *ABCB1* differences between the Vlax and non-Vlax Roma subpopulations obtained in this study clearly show that the subdivision of the Roma population is an important factor in pharmacogenetic analyses. Due to the Roma endogamy and isolation, the ADME gene-per-gene and locus-per-locus genetic specificities should be expected in this population and taken into account in pharmacological practice.

CONCLUSION

We confirmed the initial hypothesis that due to the genetic distinctiveness of the Roma population caused by their specific origin and centuries-long reproductive isolation, the prevalence of alleles in polymorphic loci ABCB1 pharmacogene differs in comparison with other populations, which is clinically relevant.

Acknowledgements

The research was funded by the Croatian Science Foundation grant (IP-2014-09-4454) to M. Pericic Salihovic. The samples were collected within the projects funded by Croatian Ministry of Science, Education and Sports (196-1962766-2747, 196-1962766-2763) to N. Smolej Narancic and B. Janicijevic, respectively, and Nutricia Research Foundation (2012-36/2013-E7) to T. Skaric-Juric.

REFERENCES

1. Veeramah KR, Novembre J. Demographic events and evolutionary forces shaping European genetic diversity. Cold Spring Harb Perspect Biol. 2014;6(9):a008516.
2. Fraser A. The Gypsies. Blackwell Publishers, Oxford, UK (1992).
3. Hancock I. We are the Romani people. UH Press, Hatfield, UK (2002).
4. Macekova S, Bernasovsky I, Gabrikova D, Bozikova A, Bernasovska J, Boronova I, Behulova R, Svickova P, Petrejcíkova E, Sotak M, Sovicova A, Carnogurska J. Association of the FTO rs9939609 polymorphism with obesity in Roma/Gypsy population. Am J Phys Anthropol. 2012;147:30-34.
5. Janicsek I, Sipeky C, Bene J, Duga B, Melegh BI, Sümegi K, Jaromi L, Magyari L, Melegh B. Significant interethnic differences in functional variants of PON1 and P2RY12 genes in Roma and Hungarian population samples. Mol Biol Rep. 2015; 42: 227-232. doi: 10.1007/s11033-014-3762-9
6. Nagy K, Fiatal S, Sándor Js, Ádány R. Distinct Penetrance of Obesity-Associated Susceptibility Alleles in the Hungarian General and Roma Populations. Obes Facts. 2017;10(5):444-457.
7. Gil-PeNa H, Coto E, Santos F, Espino M, Cea Crespo JM, Chantzopoulos G, Komianou F, Gomez J, Alonso B, Iglesias S, Treard C, Vargas-Poussou R; Renaltube Group. A new SLC12A3 founder mutation (p.Val647Met) in Gitelman's syndrome patients of Roma ancestry. Nefrología. 2017;37(4):423-428.
8. Dimishkovska M, Kotori VM, Gucev Z, Kocheva S, Polenakovic M, Plaseska-Karanfilska D. Novel Founder Mutation in FANCA Gene (c.3446_3449dupCCCT) Among Romani Patients from the Balkan Region. Balkan Med J. 2018;35(1):108-111.
9. Borst P, Oude Elferink R. Mammalian ABC Transporters in Health and Disease. Annu Rev Biochem. 2002; 71(1): 537-592.
10. Klein I, Sarkadi B, Váradi A. An inventory of the human ABC proteins. BBA - Biomembranes 1999; 1461(2): 237-262.
11. Dean M. The Human ATP-Binding Cassette (ABC) Transporter Superfamily [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2002 Nov 18. Available from: https://www.ncbi.nlm.nih.gov/books/NBK3/
12. Hodges LM, Markova SM, Chinn LW, Gow JM, Kroetz DL, Klein TE, Altman RB. Very important pharmacogene summary: ABCB1 (MDR1, P-glycoprotein). Pharmacogenet Genomics. 2011;21(3):152-61.
13. Brinkmann U, Eichelbaum M. Polymorphisms in the ABC drug transporter gene MDR1. Pharmacogenomics J. 2001;1:59-64.
14. Fromm MF. Importance of P-glycoprotein at blood-tissue barriers. Trends Pharmacol Sci. 2004;25:423-429.
15. Ambudkar SV, Dey S, Hrycyna CA, Ramachandra M, Pastan I, Gottesman MM. Biochemical, cellular, and pharmacological aspects of the multidrug transporter. Annu Rev Pharmacol Toxicol. 1999;39:361-398.
16. Raviv Y, Puri A, Blumenthal R. P-glycoprotein-overexpressing multidrug-resistant cells are resistant to infection by enveloped viruses that enter via the plasma membrane. FASEB J. 2000;14:511-515.
17. Litman T, Zeuthen T, Skovsgaard T, Stein WD. Competitive, non-competitive and cooperative interactions between substrates of P-glycoprotein as measured by its ATPase activity. Biochim Biophys Acta. 1997;1361:169-176.
18. www.pharmaadme.org
19. Abraham JE, Guo Q, Dorling L, Tyrer J, Ingle S, Hardy R, Vallier AL, Hiller L, Burns R, Jones L, Bowden SJ, Dunn JA, Poole CJ, Caldas C, Pharoah PP, Earl HM. Replication of genetic polymorphisms reported to be associated with taxane-related sensory neuropathy in patients with early breast cancer treated with Paclitaxel. Clin Cancer Res. 2014;20(9):2466-2475.
20. Zhou Z, Chen Q, Zuo D, Wang H, Hua Y, Cai Z. ABCB1 (rs1128503) polymorphism and response to chemotherapy in patients with malignant tumors-evidences from a meta-analysis. Int J Clin Exp Med. 2015;8(1):265-272.
21. Jmel H, Romdhane L, Ben Halima Y, Hechmi M, Naouali C, Dallali H, Hamdi Y, Shan J, Abid A, Jamoussi H, Trabelsi S, Chouchane L, Luiselli D, Abdelhak S, Kefi R. Pharmacogenetic landscape of Metabolic Syndrome components drug response in Tunisia and comparison with worldwide populations. PLoS One 2018;13(4):e0194842.
22. Wang D, Johnson AD, Papp AC, Kroetz DL, Sadee W. Multidrug resistance polypeptide 1 (MDR1, ABCB1) variant 3435C>T affects mRNA stability. Pharmacogenetics. 2005;15(10):693-704.
23. Mega JL, Close SL, Wiviott SD, Shen L, Walker JR, Simon T, Antman EM, Braunwald E, Sabatine MS. Genetic variants in ABCB1 and CYP2C19 and cardiovascular outcomes after treatment with clopidogrel and prasugrel in the TRITON-TIMI 38 trial: a pharmacogenetic analysis. Lancet 2010;376(9749):1312-1319.
24. Leschziner G, Zabaneh D, Pirmohamed M, Owen A, Rogers J, Coffey AJ, Balding DJ, Bentley DB, Johnson MR. Exon sequencing and high resolution haplotype analysis of ABC transporter genes implicated in drug resistance. Pharmacogenet Genomics. 2006;16:439-450.
25. Semagn K, Babu R, Hearne S, Olsen M. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. Mol Breeding. 2013;33:1-14.
26. Takezaki N, Nei M, Tamura K. POPTREE2: Software for Constructing Population Trees from Allele Frequency Data and Computing Other Population Statistics with Windows Interface. Mol Biol Evol 2010;27(4):747-752.
27. Skaric-Juric T, Tomas Z, Zajc Petranovic M, Bozina N, Smolej Narancic N, Janicijevic B, Pericic Salihovic M. Characterization of ADME genes variation in Roma and 20 populations worldwide. PloS One. 2018;13(11):e0207671.
28. https://www.medcalc.org/calc/comparison_of_proportions.php
29. Sporis D, Basic S, Bozina N, Babic T, Hajnsek S, Sertic J, Susak I, Markovic I. ABCB1 gene variants as predictors of multidrug-resistant epilepsy in Croatian population. Neurol Croat. 2011;60:63-70.
30. Sipeky C, Lakner L, Szabo M, Takacs I, Tamasi V, Polgar N, Falus A, Melegh B. Interethnic differences of CYP2C9 alleles in healthy Hungarian and Roma population samples: relationship to worldwide allelic frequencies. Blood Cells Mol Dis. 2009;43(3):239-242.
31. Weber A, Szalai R, Sipeky C, Magyari L, Melegh M, Jaromi L, Matyas P, Duga B, Kovesdi E, Hadzsiev K, Melegh B. Increased prevalence of functional minor allele variants of drug metabolizing CYP2B6 and CYP2D6 genes in Roma population samples. Pharmacol Rep. 2015;67:460-464.
32. Tomas Z, Kuhanec A, Skaric-Juric T, Zajc Petranovic M, Smolej Narancic N, Janicijevic B, Salihovic MP. Distinctiveness of the Roma population within CYP2B6 worldwide variation. Pharmacogenomics. 2017;18(17):1575-1587.
33. Nagy A, Sipeky C, Szalai R, Melegh BI, Matyas P, Ganczer A, Toth K, Melegh B. Marked differences in frequencies of statin therapy relevant SLCO1B1 variants and haplotypes between Roma and Hungarian populations. BMC Genet. 2015; 6:108.
34. Sipeky C, Csongei V, Jaromi L, Safrany E, Polgar N, Lakner L, Szabo M, Takacs I, Melegh B. Vitamin K epoxide reductase complex 1 (VKORC1) haplotypes in healthy Hungarian and Roma population samples. Pharmacogenomics. 2009;10(6):1025-1032.

35. Janicsek I, Sipeky C, Bene J, Duga B, Melegh B, Sumegi K, Jaromi L, Magyari L, Melegh B. Significant interethnic differencies in functional variants of pon1 and p2ry12 genes in Roma and Hungarian population samples. Mol Biol Repts. 2015;42:227-232.
36. Pericic Salihovic M, Baresic A, Martinovic Klaric I, Cukrov S, Barac Lauc L, Janicijevic B. The role of the Vlax Roma in shaping the European Romani maternal genetic history. Am. J. Phys. Anthropol. 2011;146(2):262-270.
37. Fung KL, Gottesman MM. A synonymous polymorphism in a common MDR1 (ABCB1) haplotype shapes protein function. Biochim Biophys Acta. 2009;1794(5):860-71. doi: 10.1016/j.bbapap.2009.02.014.
38. Haerian BS, Lim KS, Tan CT, Raymond AA, Mohamed Z. Association of ABCB1 gene polymorphisms and their haplotypes with response to antiepileptic drugs: A systematic review and meta-analysis. Pharmacogenomics. 2011;12:713-725.
39. Vivona D, Lima LT, Rodrigues AC, Bueno CT, Alcantara GK, Barros LS, DE Moraes Hungria VT, Chiattone CS, DE Lourdes Lopes Ferrari Chauffaille M, Guerra-Shinohara EM. ABCB1 haplotypes are associated with P-gp activity and affect a major molecular response in chronic myeloid leukemia patients treated with a standard dose of imatinib. Oncol Lett. 2014;7(4):1313-1319.
40. Li H, Wang B, Chang C, Wu M, Xu Y, Jiang Y The Roles of Variants in Human Multidrug Resistance (MDR1) Gene and Their Haplotypes on Antiepileptic Drugs Response: A Meta-Analysis of 57 Studies. PLoS One. 2015;10(3):e0122043.
41. Martinez-Cruz B, Mendizabal I, Harmant C, de Pablo R, Ioana M, Angelicheva D, Kouvatsi A, Makukh H, Netea MG, Pamjav H, Zalan A, Tournev I, Marushiakova E, Popov V, Bertranpetit J, Kalaydjieva L, Quintana-Murci L, Comas D. Origins, admixture and founder lineages in European Roma. Eur J Hum Genet. 2016;24(6):937-943.
42. Kalaydjieva L, Calafell F, Jobling MA, Angelicheva D, de Knijff P, Rosser ZH, Hurles ME, Underhill P, Tournev I, Marushiakova E, Popov V. Patterns of inter- and intra-group genetic diversity in the Vlax Roma as revealed by Y chromosome and mitochondrial DNA lineages. Eur J Hum Genet. 2011;9:97-104.
43. Mendizabal I, Lao O, Marigorta UM, Wollstein A, Gusmao L, Ferak V, Ioana M, Jordanova A, Kaneva R, Kouvatsi A, Kučinskas V, Makukh H, Metspalu A, Netea MG, de Pablo R, Pamjav H, Radojkovic D, Rolleston SJ, Sertic J, Macek M Jr, Comas D, Kayser M. Reconstructing the population history of European Romani from genome-wide data. Curr Biol. 2012;22:2342-2349.
44. Moorjani P, Patterson N, Loh P-R, Lipson M, Kisfali P, Melegh BI, Bonin M, Kadasi L, Rieß O, Berger B, Reich D, Melegh B. Reconstructing Roma History from Genome-Wide Data. PLoS ONE. 2013;8(3):e58633.