INVESTIGATION OF THE GENETIC VARIABILITY AND PHARMACOLOGICAL SIGNIFICANCE OF THE ORGANIC ANION TRANSPORTER PROTEINS IN HUNGARIAN AND ROMA POPULATIONS

DOCTORAL (Ph.D.) THESIS

DR. ÁGNES NAGY

Doctoral school leader

Prof. Dr. Balázs Sümegi

Supervisor

Prof. Dr. Béla Melegh

University of Pécs

Faculty of Medicine

Interdisciplinary Medical Sciences Doctoral School

Clinical Center, Department of Medical Genetics

Pécs, 2017

1. INTRODUCTION

Metabolism is the set of life-sustaining chemical transformations of different endogenous and exogenous chemical materials within the cells of living organisms, which results modified, more polar, more easily eliminated materials. In the transformation of the different substances there are many different pathways and enzymes involved, which are capable to the synthesis and degradation of different substrates.

Drugs can overpass the biological membranes in several ways. The influx and efflux of the essential materials for the organism including glucose, amino acids, inorganic materials, ions and different drugs are controlled by different transmembrane proteins. These proteins could be divided to active and passive transporters based on their behavior. Usually the influx and the efflux transporters also exist in one cell parallel. Proteins which have transport function can be separated into three huge categories:

1. Active transporters (ATP pumps)
2. Ion channel proteins
3. Transporters (Carrier proteins)

DRUG TRANSPORTER PROTEINS

ABC transporters

The most of the efflux transporters belong to the ATB-binding cassette transporter (ABC transporter) superfamily, which affects the intracellular concentration of the different materials in the cells. The necessary energy for the overpassing of the substrates through the membranes comes from the hydrolysis of the ATP and from the phosphorylation of the transport protein. This process enables the movement of the substrates through the membranes depending of their concentration gradient.

The ABC superfamily contains 49 proteins, which could be differentiated to 7 sub families based on their domain organization and phylogenetic analysis. Nowadays more than 20 ABC proteins can be associated with different diseases, and most of them have a clinical significance in the drug metabolism and drug resistance.

The first identified and most characterized ABC transporter is the multidrug resistance causing human ABCB1. The protein has wide substrate specificity and serves as a functional barrier against different drugs. Expression of the protein in endothelial cells of the blood-brain barrier results the inhibition of the drug transport into the central nervous system.

The ABCC1 protein, which is coded by the *ABCC1* gene, has been first identified in doxorubicin-resistant small cell lung cancer cell lines. The ABCC1 serves as a multispecific organic anion transporter for those drugs, like antimetabolites, antracycline, plant alkaloids and antiandrogenes.

The ATP-binding cassette transporter MRP2 protein is coded by the *ABCC2* gene. ABCC2 was first described as a canalicular multispecific organic anion transporter, but exports actively more non-conjugated substrates. MRP2 makes it easier the transport of those anti-cancer drugs, like cisplatin, vinblastine and camptothecin derivatives.

The breast cancer resistance protein (BCRP) or ABCG2 is the most known member of the G-family of the ABC transporters. By interacting with hem and other porfirins it can protect the cells and tissues from the protoporfirin accumulation. It is supposed that the ABCG2 has a significant effect for the pharmacokinetic and pharmacodinamic profile of many xenobiotics and endogenous substrates. The protein could facilitate to the multidrug resistance, since it has typical substrates, like cisplatin, camptothecin, doxorubicin, daunorubicin, etopsid, methotrexate, mitoxantrone, SN-38, topotecan and vincristine.

SLC transporters

The solute carrier (SLC) family contains the most membrane transport proteins. Most of the SLC transporters are secondary active transporters, for example the ion changers, symporters and the antiporters, where the transport occurs by different energy linked mechanisms. The human SLC transporter family includes 386 members. The different proteins can be divided into 52 sub families based on the number of their alpha helixes (10-14THM) and their biological functions.

SLC proteins regulate the transport of those substrates, like the inorganic ions, nucleotides, amino acids, neurotransmitters, glucose, purines, fatty acids and drug molecules.

SLC mutations or the genetic variants of the members could play a role in the development of autism, diabetes, cancer, psychiatric disorders and neurodevelopmental disorders. Based on these facts SLC proteins act as very important therapeutic targets.

The glutamate transporters (GLT) belong to the SLC1 family and keep the extracellular glutamine concentration below the excitotoxic level. Pathogenesis of the amyotrophic lateral sclerosis, the Alzheimer-disease and the autism could be associated with the SLCA2 member of the family. SLC3A plays a role in the pathogenesis of the schizophrenia. In ischemia the neuronal glutamate transporter SLC1A1 probably serves as a reverse glutamate transporter.

Because of this, glutamate transporter specific inhibitors could be potential therapeutic possibilities for the prevention of the excitotoxicity during ischemic environment.

The SLC2A9, which belongs to the SLC2 family was first considered as a glucose or fructose transporter.

Pharmaceutically the SLC6 is the most well investigated and used SLC family. The proteins regarding this family transport serotonin, dopamine, noradrenalin, gamma amino acid, taurine and creatinine. The proteins of this family can be associated to attention-deficit hyperactivity disorder, X-linked mental retardation, Tourette-syndrome, schizophrenia, Parkinson-disease, autism, depression, anxiety, obsessive compulsive disorder and posttraumatic stress disease.

The SLC13 family contains the Na^+-linked di- and tricarbonate/sulfate transporters. SLC13A2 and SLC13A3 proteins have highlighted clinical importance. Latter can be associated with aciduria and Canavan-disease.

Vesicular monoamine transporters (VMAT) are responsible for the transport of the monoamines to synaptic vesicles. Genetic association tests detected that *VMAT1* variants can be related to anxiety disorders, schizophrenia and bipolar disorders. Increased VMAT2 activity in case of Parkinson-disease could create a new therapeutic target or could optimize the prognosis.

The SLC21 (organic anion transporter), the SLC22 (organic cation/anion/zwitterion transporter) and the SLC47 (multidrug and toxin extrusion (MATE) transporter) proteins are expressed in the liver, kidneys and blood-brain barrier, where mediates the absorption, distribution, metabolism and excretion of the drugs.

THE GENETICS OF THE *SLCO* GENES

The organic anion transporter polypeptides (OATPs) coded by the *SLCO* genes are membrane-linked pharmaceutical agent transporting proteins, which facilitate the drug uptake to the cells. SLCO1 member of the protein family are responsible for the transport of the drug substances, while SLCO3, SLCO5 and SLCO6 have a role in the transport of the organic anions. SLCO2 is responsible for the prostaglandin and steroid sulfate transport, while SLCO4 acts as a thyroid hormone transporter.

SLCO1B1 gene

The OATP1B1 coded by the solute carrier organic anion transporter family 1B1 gene (*SLCO1B1*) is responsible for the uptake of different endogenous substrates (for example bile acids), xenobiotics, different drug substances (for example: statins, antibiotics, angiotensin-converting enzyme (ACE) inhibitors) to the cells. The *SLCO1B1* gene has 190 known variants; its minor frequency is bigger than 5%.

OATP1B1 has a significant role in the pharmacokinetics of the statins. Statins are 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors, and are widely used to reduce the risk of the cardiovascular diseases. In the last few years an association has been found between the *SLCO1B1* variants and the simvastatin-induced myopathy, referring to the fact that OATP1B1 plays a role in the simvastatin transport.

Among the 190 known polymorphisms of the *SLCO1B1* gene, the most characterized variants are the rs2306283 (c.388A>G, p.Asn130Asp) and the rs4149056 (c.521T>C, p.Val174Ala) ones.

The c.388A>G SNP of the *SLCO1B1* results the increased activity of the OATP1B1 protein, and a decreased plasma statin concentration. The c.521T>C polymorphism is associated to a reduced protein activity and an increased plasma statin concentration. The two SNPs together determine four haplotypes: SLCO1B1*1A (c.388A - c.521T, wild type), SLCO1B1*1B (c.388G - c.521T), SLCO1B1*5 (c.388A - c.521C) and SLCO1B1*15 (c.388G - c.521C). From these the SLCO1B1*1B is the most common, which is followed by the SLCO1B1*15 and the SLCO1B1*5.

In 60% of the cases having the non-coding rs4363657 (c.1498-1331T>C) polymorphism of the *SLCO1B1* gene an association could be detected between the C variant and the statin-induced myopathy.

SLCO1B3 gene

The OATP1B3 (organic anion-transporting polypeptide 1B3) is an important transmembrane protein, which conveys the Na^+-independent uptake of endogenous and exogenous substances. The OATP1B3 influx protein plays an essential role in the pharmacokinetics of the digoxin and statins used for the treatment of heart failure.

The two most commonly investigated missense variants of the *SLCO1B3* gene are the c.334T>G (rs4149117, p.Ser112Ala) and the c.699G>A (rs7311358, p.Met233Ile) polymorphisms of the exons 4 and 7.

The c.344T>G (p.Ser112Ala) polymorphism in carriers of a kidney transplanted and mycophenolate mofetil treated cohort have an altered pharmacokinetic effect.

Based on earlier investigations the intronic c.1683-5676A>G (rs11045585) variant of the *SLCO1B3* gene could be associated by decreased docetaxel clearance values. In addition the presence of this SNP showed the most significant association between the docetaxel-induced leukopenia.

STATIN TERAPHY

Cholesterol is an essential sterol of the human organization, which has a very important role in the build-up of the cell membranes and this is the originating substance of many hormone synthesis. Reduction of the cholesterol level reduces the risk of the arteriosclerosis and the juvenile Coronary Vascular Disease (CVD).

Statins are HMG-CoA reductase (HMGCR) inhibitors. The HMGCR plays a central role in the synthesis of the cholesterol. The higher level of cholesterol means the higher risk for cardiovascular diseases. Many controlled clinical examinations verified that the statin therapy decreases the risk for the myocardial vascular events, and the frequency of the occurrence of the stroke events.

From five huge clinical investigations two confirmed that more intensive statin therapy could be associated with the reduction of the number of coronary disease cases. The only obstruction of the widely usage of the high dosage statin therapy is the safety judgment of the drug, and the enumeration of the possible side effects. The side effects of the statins at higher dosage usage could occur with a higher risk, however some genetic polymorphisms could have a great influence for the statin metabolism.

SIDE EFFECTS OF STATINS

Myopathy

Limiting factor of the statin therapy could be the dosage dependent and genetic variability-dependent myopathy side effect. The clinical picture of the statin myopathy most commonly includes symmetrical lower limb weakness and muscle pain. Incidence could be 1.5%. In case of higher statin dosage there is a higher chance for the development of statin-induced myopathy.

Rhabdomyolysis

Rhabdomyolysis is a musculoskeletal-related, potentially fatal condition, in which the skeletal muscle cells die and the myoglobin first enters to the blood, and finally eliminates through the kidneys. Clinical symptoms are the weakness, muscle pain, but symptom free forms also can be observed. The urine becomes darker. Laboratory symptoms: CK, SGOT, SGPT, LDH increases. In 1% of cases occasionally myopathy could occur as a serious side effect.

Statins and the diabetes

The statin usage in diabetic cases without coronary disease shows the same decrease of risk like in other cohorts in case of coronary disease occasions. On the other side more clinical investigations including 91140 cases resulted, that statins could increase the risk of the development of diabetes. In total the benefits of statin therapy is essential in this group of patients.

Other side effects of statins

Usage of statins could cause other side effects, for example the increase of the liver function values (primarily GOT and GPT), pancreatitis, hepatitis (including the chronic active hepatitis), cholestasis, fatty liver, cirrhosis, fulminant hepatitis, hepatoma, anorexia, nausea, vomiting and memory disturbances.

Genetic variants affecting the statin therapy

It is well known, that some polymorphisms of the *SLCO1B1* transporter gene could be associated with a higher risk for the simvastatin induced statin myopathy, thus these variant genotypes result higher statin concentration. This means, that there is an increased risk in case of high dosage statin usage. Rhabdomyolysis could occur in more cases over 80 mg simvastatin treatment. The incidence of the rhabdomyolysis could occur in about 1.9/100.000. In 60% of these cases the rs4149056 genetic variant stands in the background of the development of the rhabdomyolysis.

THE HUNGARIAN AND THE ROMA POPULATIONS

The Hungarian population

The Hungarians are unique among the other surrounding populations regarding to their origination. The Hungarian government was established 1100 years ago. The early Hungarians settled down in the Carpathian-basin at the end of the 9th century after two thousand years migration, leaving behind the Ural Mountains. Thousands of years before the arrive of the Hungarians Dacians, Romans, Sarmatas, Gothic people, Hun people Avars and Slavs were lived already in this place. At the time of the arrival to the Carpathian-basin most of the indigenous people had Slavic origination.

Mitochondrial DNA examinations, chromosome Y binary marker examinations and array based SNP examinations were used for the investigation of the populations living in the Carpathian-basin about 1100 years ago. The aim of the investigations was to explore the genetic origin of the ancient Hungarian populations.

Mitochondrial sequence of 27 ancient Hungarian samples, 101 recent Hungarian samples and 76 Hungarian samples from Transylvania was compared to 7752 samples from 57 European and Asian population, including the Finno-Ugrian populations also. The results showed that the ancient Hungarian population from the 10-11th century was genetically heterogeneous, and a small Asian effect also could be detected in the ancestry population.

Paternal inheritance was also investigated, which results better geographical distribution, than the maternal one. Totally 22 biallelic polymorphism was detected in the non-recombinant region of the human chromosome Y of 100 modern Hungarians and 97 Szeklers. The results were compared to other European populations and pools of the chromosome Y were analyzed in phylogeographic context. A special new chromosome Y base change (T>C) (95% confidence interval, 3140-6200 years) is very valuable marker in the investigation of the Finno-Ugrian populations. The C allele of this polymorphism is widespread in all populations in the Ural language group, except the modern Hungarian speaking populations. In these populations this polymorphism is totally missing or very rare. Among the modern samples only 1 Szekler sample carried this C allele, while from the 4 ancient Hungarian bone samples only 2 contained it. Based on these results the ancient Hungarians had a Siberian filiation, which disappeared later.

The lactase non-persistence (hypolactasia) shows autosomal recessive inheritance pattern. The prevalence of the adult lactase non-persistence is 3-70% in the Caucasian populations of Europe. The disease in North-Europe is rare, but in south and in east it is more frequent. The

Asian populations show almost 100% prevalence. Recently the T>C SNP of the *LTC* gene was associated to the lactase non-persistence. Further examinations resulted, that the C/T-13910 polymorphism has a role in the mediation of the lactase gene expression. The prevalence of the different C/T-13910 lactase genotypes was examined in the recent Hungarian populations using random sampling. The frequency of the T allele was 37.8%, while the C allele showed 62.2% frequency. The frequency of the C allele in the Hungarian populations was lower, than in the Swedish and Finnish populations (81%), but Hungarians showed higher values, than in the French (43.1%) and Northern-Italian (35.7%) populations. The Hungarian C allele frequency showed the same value like the Portuguese population (62%). On the other side those populations, which lived near the Hungarians in their Siberian home showed the same high C allele frequency: Northern-Manysiak (71%), Nyenyecs (78%), Komi-Permjacs (50%) and Udmurtok (59%).

Roma population

Based on written historical evidences and results of linguistic and population genetics examinations, the Roma population could be originated from the North-West Indian Punjab, Rajasthan and Gujarat states. Their migration was started in the 11th century. They leaved North-West India and arrived to Europe to the 13th century through Iran. From the end of the 14th century the Roma population reached all European countries.

Y haplogroup investigations showed that 47.3% of the Roma males carries the Y chromosome H-M82 haplogroup, which could be detected rarely outside the Indian subcontinent. The mitochondrial M-haplogroup is the most common in the Indian people, and could be detected rarely outside South-Asia, but approximately 30% of Roma people have it. Detailed analysis of Polish Roma samples also indicated the Indian-specific M5 haplogroup. Besides the Y haplogroup analyses, whole genome based SNP array and Next Generation Sequencing results also confirmed the assumptions about the Indian origin of the Roma people.

Arriving to Europe, due to a bottleneck effect the Roma population size reduced, generating a small number founder population. Based on this process and because of the closed genetic system of the Roma communities, Roma populations have a unique genetic profile.

Hungary is a heterogeneous population, containing more ethnic minorities. Roma minority creates the biggest group of the Hungarian ethnicities. Based on a 2008 year data the number of the Roma people in Hungary was about 600,000-800,000.

Based on their history and language the Hungarian Roma people could be divided three big groups: Romungró, Oláh and Beás.

Previous international studies showed that some special rare inherited diseases could be associated with the Roma populations, including inherited motor and sensory neuropathy, congenital cataract, dysmorphic face, neuropathy syndrome, congenital myasthenia syndrome, limb-girdle muscular dystrophy, lack of galactokinase and polycystic kidney disease. As a result of the closed genetic system an increased autozygosity occurred in Roma colonies, which contributed to the accumulation of the autosomal recessive multiplex disorders.

2. AIMS OF THE STUDY

During the investigation our aim was the genetic examination of the 5 polymorphisms of the organic anion transporter coding *SLCO1B1* and *SLCO1B3* genes, including the determination of their frequency and distribution in Roma and Hungarian samples.

We have investigated the following polymorphisms:

1. *SLCO1B1* gene c.388A>G (rs2306283) polymorphism
2. *SLCO1B1* gene c.521T>C (rs4149056) polymorphism
3. *SLCO1B1* gene c.1498-1331T>C (rs4363657) polymorphism
4. *SLCO1B3* gene c.334T>G (rs4149117) polymorphism
5. *SLCO1B3* gene c.1683-5676A>G (rs11045585) polymorphism

Further aim was the investigation of the linkage of the three *SLCO1B1* and the two *SLCO1B3* polymorphisms, determination of the created haplogroups, and their frequency in Roma and Hungarian populations.

3. MATERIALS AND METHODS

OBSERVED POPULATIONS

The Hungarian and Roma samples used for our investigations were collected from healthy Hungarian Roma and Hungarian persons. The DNA samples were from the Central Biobank of the University of Pécs, which is the member of the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI).

Our investigations included the examination of 5 functionally significant polymorphisms of the *SLCO1B1* and *SLCO1B3* genes.

In case of the *SLCO1B1* rs2306283 (c.388A>G), rs4149056 (c.521T >C) and rs4363657 (c.1498-1331T>C) polymorphisms 470 Roma (170 man and 300 woman; average age 39±16 years) and 442 Hungarian (183 man and 259 woman; average age 45±10 years) persons were investigated. In case of *SLCO1B3* rs4149117 (c.334T>G) and rs11045585 (c.1683-5676A>G) polymorphisms 467 Roma (172 man and 295 woman, average age 39±15 years) and 448 Hungarian (204 man and 244 woman; average age 45±11 years) persons were investigated.

MOLECULAR BIOLOGY METHODS

DNA isolation

DNA isolation was performed from EDTA anti-coagulated blood samples.

Polymerase chain reaction

The starting point of the DNA-analysis was the amplification by PCR with a 50µl volume. PCR was performed using home-designed specific primer pairs (forward and reverse), dNTP, Taq polymerase, buffer and genomic DNA-template.

Further analysis of the PCR results was made by gel-electrophoresis with ethidium-bromide staining and UV lightning.

Direct sequencing

The verification of the specificity of the home-designed PCR-RFLP methods and our results were performed by Sanger bidirectional sequencing on random samples. For the sequencing process BigDye Terminator v.1.1 cycle sequencing kit and ABI 3500 Genetic Analyser were used.

STATISTICAL ANALYSIS

The associations between the populations and the examined genetic variants were investigated by $\chi 2$-test and regression analysis using the SPSS 20.0 software package. The significance level was $p<0.05$. The haplotype analysis was performed by Phase 2.1. Software, while the linkage analysis was done by Haploview 3.3 software.

4. RESULTS

SLCO1B1 GENE

Determination of genotype- and allele frequency

Genotype- and phenotype allele frequencies of the *SLCO1B1* 388A>G, 521T>C and 89595T>C polymorphisms in Roma and Hungarian populations can bee seen in Table 1.. Based on the investigation of the *SLCO1B1* rs2306283 polymorphism, we can conclude that there is a statistical significant difference between the frequency of the variant and wild type homozygous genotypes between the Roma and Hungarian populations. The frequency of occurrence of the *SLCO1B1* 388AA wild type genotype was 24.5% in Roma samples, while it was 45.5% in Hungarian samples. The genotype frequency of the 388GG homozygous variant was 33.4% in the Roma and 17.9% in the Hungarian populations. Significant difference was observed in the frequency of the AG+GG (75.5% vs. 54.5%) variant carriers and in the *SLCO1B1* 388G allele frequency (54.5% vs. 36.2%, p<0,001) in the two investigated populations.

In case of the *SLCO1B1* rs4149056 polymorphism, we can conclude that there is a statistical significant difference between the frequency of the *SLCO1B1* 521TT wild genotype between the two populations. The Roma population resulted 67.0%, while the Hungarian showed 65.2% (p=0.05). In contrary there was a difference between the frequency of the homozygous *SLCO1B1* 521CC genotype (1.49% vs. 2.94%) and the *SLCO1B1* 521C variant allele (17.2% vs. 18.9%), but it was not significant.

The investigation of the intronic *SLCO1B1* c.1498-1331T>C rs4363657 polymorphism gave the same results in the Roma and in the Hungarian populations. The frequency of the homozygous CC genotype and the *SLCO1B1* 1498-1331C allele was slightly increased in Hungarian samples comparing to the Roma samples (3.6% vs. 2.6% and 19.6% vs. 18.5%). Besides, the *SLCO1B1* 1498-1331 TC heterozygous genotype showed the same frequency in the two observed populations (31.9%).

Haplotype analysis

The investigation of the 3 examined variant of the *SLCO1B1* gene resulted 8 main haplotypes (ht). The most common haplotype was the ht8 (GTT) in both population samples, with a 43.6% incidence in Roma samples, and a 59% incidence in Hungarians. The ht6 haplotype could not be observed in the Roma samples, Hungarians also showed only 0.18%

frequency. The haplotype analysis resulted statistically significant differences in the frequency of the ht4 (ATT, 37.2% vs 20.8%), ht5 (GCC, 1.15% vs. 3.62%) and ht8 (GTT, 43.6% vs. 59.1%) haplotypes. Between these three value pairs the significance values were always $p<0.01$.

Linkage analysis

Based on the linkage analysis of the investigated *SLCO1B1* rs2306283, rs4149056 and rs4363657 variant alleles, we can conclude that, there is a totally linkage disequilibrium between the rs4149056 and rs4363657 polymorphisms both in Roma (LD=95) and in Hungarian (LD=96) populations. Furthermore a strong linkage could be observed in Roma populations between the *SLCO1B1* rs2306283 and rs4149056 SNPs (LD=86).

SLCOIB3 GENE

Determination of genotype- and allele frequency

Based on the investigation of the *SLCO1B3* c.334T>G (rs4149117) polymorphism the frequency of the *SLCO1B3* 334GG homozygous genotype was significant higher in Roma samples, than in Hungarian ones (41.54% vs. 8.04%, $p<0.001$). Comparing the Roma and Hungarian samples further significant difference was observed in the frequency of the *SLCO1B3* 334G variant allele (70.56% vs. 52.23%, p=0,001). (Table 2.)

In case of the intronic *SLCO1B3* c.1683-5676A>G (rs11045585) variant, significant difference was noticed in the frequency of the 1683-5676G variant in the investigated populations (3.43% vs. 15.07%, $p<0.001$). The homozygous variant *SLCO1B3* 1683-5676GG genotype was significant higher in Hungarians, than in Roma populations (2.01% vs. 0.43%, p=0.028). (Table 2.)

Linkage analysis

Linkage disequilibrium analysis was performed for the investigation of the linkage of the coding *SLCO1B3* c.334T>G (rs4149117) and the intronic c.1683-5676A>G (rs11045585) polymorphisms. The LD values (|D'|x100) in the Roma and Hungarian populations were 80 and 90, which presumes strong linkage.

5. DISCUSSION AND CONCLUSIONS

The present study includes the genetic investigation of the rs2306283, rs4149056, rs4363657, rs4149117 and rs11045585 variants of the *SLCO1B1* and *SLCO1B3* organic anion transporter coding genes in Roma and Hungarian populations, which plays a role in the transporter mediated drug uptake. The genetic investigations included the determination of the frequency and incidence of the investigated variants and genotypes in healthy samples. Single nucleotide polymorphisms of the *SLCO1B3* gene could explain in different degrees the variability of the pharmacokinetics of the anti-cancer and immunosuppressant drugs in patients thanks to the modified transporter activity.

SLCO1B1

Based on our investigations the c.388G allele of the *SLCO1B1* rs2306283 SNP could be regarded as a minor allele in the Hungarian samples similar than in the other Caucasian populations. In Roma samples - like in the Singapore Roma samples - the 388A allele seemed to be the minor allele. (Table 3.)

SLCO1B1 c.521T>C (rs4149056) SNP is a common polymorphism in different populations, including Caucasians (8-20%), Chinese (16%) and Japanese (10-16%). The minor allele frequency of the *SCLO1B1* 521C variant is approximately 3 times higher in Roma populations than in other Indian ones (17.2% vs. 6.5%). (Table 3.)

The 18.9% allele frequency of the *SLCO1B1* 521C in Hungarian samples is similar high, than in other Caucasian populations. SLCO1B1*15 haplotype is generated by the 521C variant, which could be associated to the rifampin-induced liver damage. Furthermore the incidence of this missense polymorphism increases the systematic exposition in case of simvastatin therapy resulting an increased risk to the simvastatin-induced myopathy.

In Hungarian cases the allele frequency of the 89595C was slightly higher, than in Roma samples or in other European populations. Surprisingly the frequency of the *SLCO1B1* 89595C allele was three times higher in Roma samples than in other Indian populations (Guajarati), but resulted similar high values like previous African observations. Comparing our Roma and Hungarian allele frequency values to non-HapMap results of other researchers regarding to the *SLCO1B1* rs4363657 SNP, it is obvious that our results are the same as the Caucasian ones. Carrying the 89595C intronic variant contributes to a higher risk for simvastatin-induced myopathy like the 521T>C SNP. (Table 4.)

Based on the haplotype analysis of the *SLCO1B1* rs2306283, rs4149056 and r4363657 SNPs, we could conclude that the most common haplotype is the ht8 (rs2306283**G**/rs4149056**T**/rs4363657**T**) both in Roma and Hungarian cases. It was followed by the ht4 (rs2306283**A**/rs4149056**T**/rs4363657**T**) and ht1 (rs2306283**A**/rs4149056**C**/rs4363657**C**) constellations.

Ht6 haplotype (rs2306283**G**/rs4149056**C**/rs4363657**T**) could be detected only with a 0.18% frequency in Hungarian samples, while Roma samples did not carry this haplotype. The h2 haplotype (rs2306283**A**/rs4149056**C**/rs4363657**T**), which represents the 521T>C variant could be characterized by decreased transporter activity and has two times higher incidence in Roma samples, than in Hungarians.

Based on the linkage disequilibrium analysis of the *SLCO1B1* rs2306283, rs4149056 and r4363657 variants we could observe that there is an absolute linkage between the rs4149056 and rs4363657 SNPs both in Roma and Hungarian populations (LD=95 vs. LD=96).

SLCO1B3

Investigation of the c.334T>G and c.1683-5676A>G polymorphisms of the *SLCO1B3* gene resulted significant difference in the frequency of the variant alleles and homozygous variant genotypes both in Roma and Hungarian samples.

344GG homozygous genotype of the *SLCO1B3* c.334T>G could be observed more than five times higher in Roma samples compared to the Hungarians. The allele frequency of the *SLCO1B3* 334G was also significantly higher in the Roma group. Pharmacokinetic effects of the *SLCO1B3* c.334T>G polymorphism is widely investigated, but the literature data is controversial. Based on Miura findings the *SLCO1B3* 334GG genotype could be associated with the increased AUC values of the mycophenolate in patients having kidney transplantation. On the other side Picard considers that the 344T allele could stand in the background of the higher AUC value of the mycophenolate. Furthermore Bouamar findings did not showed any significant association between the *SLCO1B3* polymorphisms and the drug exposition.

Based on our results the frequency of the variant allele and the GG homozygote variant genotype of the intronic *SLCO1B3* c.1683-5676A>G polymorphism has proved five time higher in Hungarians. Consequently this increased frequency in Hungarians could be associated with decreased OATP1B3 function and a potentially modified drug therapy efficiency.

Results of the investigated *SLCO1B3* c.1683-5676A>G and c.334T>G polymorphisms compared to the data of the HapMap project showed that the frequency of the intronic *SLCO1B3* 1683-5676G allele in Roma people was similar low like in other Indian populations (Gujarati). (Table 5.) Otherwise the frequency of the *SLCO1B3* 334G allele in Roma samples was lower than in the Gujarati Indian samples (70.6% vs. 94.1%). It was similar to the Chinese and Japanese results.

Hungarian samples showed similar *SLCO1B3* 1683-5676G allele frequency to other European allele frequency values (15.1% vs. 14.7%), but our results regarding to the *SLCO1B3* 334G allele frequency showed significantly lower than Italian and other population's frequencies.

LD analysis of the populations resulted strong linkage between the two investigated SNPs, but Hungarian samples showed stronger linkage than Romas (LD=90 vs. LD=80).

6. SUMMARY

- The 388 G allele and the 388GG and AG+GG genotypes of the *SLCO1B1* c.388A>G SNP have a significant higher frequency in the Roma population.
- In Hungarian populations the *SLCO1B1* c.521C variant allele and the CC homozygous genotype showed higher frequency.
- In case of the intronic *SLCO1B1* c.1498-1331T>C SNP a slightly higher frequency of the 1498-1331CC genotype and the C variant have been observed in Hungarian samples, than in the Roma population.
- Constellation of the observed rs4363657, rs2306283 and rs4149056 polymorphisms generated 8 different haplotypes in Hungarians and 7 haplotypes in Roma samples.
- The most common haplotype in the Hungarian and Roma populations was the ht8 (GTT).
- The haplotype analysis resulted significantly increased frequency of the ht4 (ATT) haplotype in Roma samples and of the ht5 (GCC) and ht8 (GTT) haplotypes in Hungarian ones.
- Strong linkage disequilibrium could be observed between the *SLCO1B1* rs4149056 and rs4363657 polymorphisms both in Roma and Hungarian samples.
- The frequency of the *SLCO1B3* rs4149117 GG homozygous genotype and the 334G allele was significant higher in the Roma populations, than in the Hungarian.
- The *SLCO1B3* rs11045585 variant frequency was significant higher in Hungarian samples.
- The homozygous variant *SLCO1B3* rs11045585 GG genotype showed higher frequency in Hungarians, than in Roma people.
- LD values of the *SLCO1B3* rs4149117 and rs11045585 variants assumed a strong linkage.

7. PUBLICATIONS

PUBLICATIONS SUPPORTING THE DISSERTATION

Nagy A, Szalai R, Magyari L, Bene J, Toth K, Melegh B.
Extreme differences in SLCO1B3 functional polymorphisms in Roma and Hungarian populations. Environ Toxicol Pharmacol. 2015 May;39(3).
(IF:2.084)

Nagy A, Sipeky Cs, Szalai R, Melegh B I, 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 Genetics. 2015.
(IF:2.40)

OTHER PUBLICATIONS

1. Bock I, Melegh B, **Nagy A**, Losonczy H, Csete B, Schroder W, Kardos M, Istvan L, Jager R, Toth AM, Toth A, Falko H, Mozsik G.
Molecular biologic study and the factor VIII gene in hemophilia A. Orv Hetil. 1996 Nov 17;137(46):2573-5.

2. **Nagy A**, Melegh B, Losonczy H.
Study of the Leiden mutation (factor VQ 506), the most frequent cause of thrombophilia, in 116 thrombosis patients. Orv Hetil. 1997 Nov 2;138(44):2797-800. Hungarian.

3. Melegh B, Stankovics J, Kis A, **Nagy A**, Storcz J, Losonczy H, Mehes K.
Increased prevalence of factor V Leiden mutation in neonatal intracranial haemorrhage. Eur J Pediatr. 1998 Mar;157(3):261.
IF: 1.050

4. Stankovics J, Melegh B, **Nagy A**, Kis A, Molnar J, Losonczy H, Schuler A, Kosztolanyi G. **Incidence of factor V G1681A (Leiden) mutation in samplings from the Hungarian population**. Orv Hetil. 1998 May 10;139(19):1161-3. Hungarian.

5. Stankovics J, **Nagy A**, Mehes K, Melegh B.

Umbilical venous catheterization and development of Banti syndrome: the possible role of the factor V Leiden mutation. Eur J Pediatr. 1998 Aug;157(8):696.

IF: 1.050

6. David M, Losonczy H, **Nagy A**, Kutscher G, Meyer M.

Screening methods in genetic diagnosis of hereditary protein C deficiency. Orv Hetil. 1999 Jan 17;140(3):125-32. Review. Hungarian.

7. Gasztonyi B, Par A, Szomor A, Battyany I, **Nagy A**, Kereskai L, Losonczy H, Mozsik G. **Hepatitis C virus infection associated with B-cell non-Hodgkin's lymphoma in Hungarian patients.** Br J Haematol. 2000 Aug;110(2):497-8.

IF: 3.068

8. David M, Losonczy H, Sas G, **Nagy A**, Kutscher G, Meyer M.

Identification of mutations in 15 Hungarian families with hereditary protein C deficiency. Br J Haematol. 2000 Oct;111(1):129-35.

IF: 3.068

9. Szomor A, Molnár L, **Nagy A,** David M, Alizadeh H, Kecskes M, Vidra T, Kereskai L, Pajor L, Losonczy H.

Treatment of chronic myeloid leukemia with interferon-alpha. Orv Hetil. 2000 Nov 26;141(48):2601-4. Hungarian.

10. Gasztonyi B, Par A, Szomor A, **Nagy A**, Kereskai L, Losonczy H, Pajor L, Horanyi M, Mozsik G.

Hepatitis C virus infection and B-cell non-Hodgkin's lymphoma. Orv Hetil. 2000 Dec 3;141(49):2649-51. Hungarian.

11. Kecskes M, **Nagy A,** Vidra T, Kispal G, Radvanyi G, Vezendi K, Hajnal L, Kellner R, Losonczy H.

Screening for carrier state of Haemophilia B using indirect genomic detection. Orv Hetil. 2001 Feb 18;142(7):341-4. Hungarian.

12. Komlosi K, Havasi V, Bene J, Ghosh M, Szolnoki Z, Melegh G, **Nagy A**, Stankovics J, Csaszar A, Papp E, Gasztonyi B, Toth K, Mozsik G, Romics L, ten Cate H, Smits P, Mehes K, Kosztolanyi G, Melegh B.
Search for factor V Arg306 Cambridge and Hong Kong mutations in mixed Hungarian population samples. Acta Haematol.2003;110(4):220-2.
IF: 1.874

13. Molnar L, **Nagy A**, David M, Szomor A, Mehes G, Kovacs G, Losonczy H.
Results of imatinib therapy in late-stage chronic myeloid leukemia after treatment with interferon-alpha. Orv Hetil. 2004 Apr 25;145(17):901-7. Hungarian

14. Toth O, David M, Habon T, **Nagy A**, Keszthelyi Z, Kovacs N, Losonczy H.
Type I antithrombin deficiency as a cause of arterial and venous thrombosis in a family with severe thrombophilia. Orv. Hetil. 2005 Oct 9;146(41):2121-5. Review. Hungarian.

15. Szilagyi A, **Nagy A**, Tamas P, Vizer M, Szabo I, Losonczy H.
Two successful pregnancies following eight miscarriages in a patient with antithrombin deficiency. Gynecol Obstet Invest. 2006;61(2):111-4. Epub 2005 Oct 21.
IF: 0.874

16. Toth O, Szabo C, Kecskes M, Poto L, **Nagy A**, Losonczy H.
In vitro effect of the potent poly(ADP-ribose) polymerase (PARP) inhibitor INO-1001 alone and in combination with aspirin, eptifibatide, tirofiban, enoxaparin or alteplase on haemostatic parameters. Life Sci. 2006 Jun 20;79(4):317-23.
IF: 2.389

17. Szendrei T, Magyarlaki T, Kovacs G, **Nagy A**, Szomor A, Molnar L, David M, Tokes-Fuzesi M, Rideg O, Poto L, Pajor L, Kajtar B, Losonczy H.
Multidrug resistance in chronic lymphocytic leukemia.
Orv Hetil. 2008 Jan 27;149(4):161-7. Hungarian.

18. Gerlinger I, Torok L, **Nagy A**, Patzko A, Losonczy H, Pytel J.
Frequency of coagulopathies in cases with post-tonsillectomy bleeding. Orv Hetil. 2008 Mar 9;149(10):441-6. Hungarian.

19. Molnar TF, Benko I, Szanto Z, **Nagy A**, Horvath OP.
Complications after ultrasonic lung parenchyma biopsy: a strong note for caution. Surg Endosc. 2008 Mar;22(3):679-82.
IF: 3.231

20. **Nagy A**, Losonczy H, Toth O, Kosztolányi Sz, Miko A, Mozes R, David M.
Masszív vérzés kezelése magas titerű inhibitoros haemophiliás betegeknél, szekvenciális áthidaló kezeléssel Hematológia- 44:(3-4) pp. 148-151. (2011)

21. Cziraki A, Ajtay Z, **Nagy A**, Marton L, Verzar Z, Szabados S.
Early post-operative thrombosis of the prosthetic mitral valve in patient with heparin-induced thrombocytopenia. J Cardiothorac Surg. 2012 Mar 13;7:23.
IF: 0.900

22. Labadi A, **Nagy A**, Szomor A, Miseta A, Kovacs GL.
Factitious hyperkalemia in hematologic disorders. Scand J Clin Lab Invest. 2017 Feb;77(1):66-72.
IF: 1.11

Impact factor of publications which have served as a base for the thesis: **4.484**

Impact factor of other published articles: **18.614**

Cumulative impact factor: 23.098

8. Acknowledgements

The research which served as a base of the thesis was carried out in the University of Pécs, Department of Medical Genetics at the Multidisciplinary Doctoral School in theme Human Genetics. First of all, I would like to thank especially to my supervisor, Professor Dr. Béla Melegh, who allowed me to take part in the research, and supported me. I also thank him for his useful advices during my work, for his leadership and help. His professional guidance was essential for the publication of the articles in international papers.

I would like to thank to Professor Dr. Kálmán Tóth, director of the 1st Department of Internal Medicine, Clinical Centre, University of Pécs for his professional help and support.

I am very grateful to the current and former employees of the Department of Medical Genetics, especially to Renáta Szalai, Dr. Judit Dr. Berenténé Bene, Dr. Erzsébet Kövesdi and Petra Mátyás for their professional support.

I want to say a specific thank to all of the assistants.

Furthermore I am also thankful to the Hungarian contributors for the collection of the investigated samples.

Table 1. Genotype- and phenotype frequency of the investigated *SLCO1B1* polymorphisms in Roma and Hungarian populations

Polymorphism	rs	Genotype	Genotype frequency	
			Roma n=470 (%)	**Hungarian n=442 (%)**
c.388A>G	rs2306283	AA	115 (24.5)	201 (45.5)
		AG	198 (42.1)	162 (36.6)
		GG	157 (33.4)[x]	79 (17.9)
		G allele frequency	54.5%[x]	36.2%
c.521T>C	rs4149056	TT	315 (67.0)[y]	288 (65.2)
		TC	148 (31.5)	141(31.9)
		CC	7 (1.5)	13 (2.9)
		C allele frequency	17.2%	18.9%
c.1498-1331T>C	rs4363657	TT	308 (65.5)	285 (64.5)
		TC	150 (31.9)	141 (31.9)
		CC	12 (2.6)	16 (3.6)
		C allele frequency	18.5%	19.6%

[x]p<0,001

[y] p=0,05

Table 2. Genotype- and allele frequency of the investigated *SLCO1B3* polymorphisms in Roma and Hungarian populations

Polymorphism	rs	Genotype	Genotype frequency	
			Roma n=467 (%)	**Hungarian n=448 (%)**
c.1683-5676A>G	rs11045585	AA	437 (93.57)	322 (71.87)
		AG	28 (6.00)	117 (26.12)
		GG	2 (0.43)***	9 (2.01)
		G allele frequency	3.43%*	15.07%
c.334T>G	rs4149117	TT	2 (0.43)	16 (3.57)
		TG	271 (58.03)	396 (88.39)
		GG	194 (41.54)*	36 (8.04)
		G allele frequency	70.56%**	52.23%

*p<0,001

**p=0,001

***p=0,028

Table 3. Genotype- and phenotype frequencies of the *SLCO1B1* G388A and T521C polymorphisms in different populations

Population	n	G388A %					T521C %					Ref.
		AA[1]	AG	GG[1]	AG+GG[1]	G allele[1]	TT[2]	TC	CC	TC+CC	C allele	
Roma	470	24.5	42.1	33.4	75.5	54.5	67.0	31.5	1.49	33.0	17.2	
Hungarian	442	45.5	36.6	17.9	54.5	36.2	65.2	31.9	2.94	34.8	18.9	
Finnish	468	29.3	49.2	21.6	70.8	46.2	63.9	31.8	4.30	36.1	20.2	[170]
Indian (North)	270	31.9	46.7	21.4	68.1	45.0	-	-	-	-	-	[171]
Indian (Singapore)	100	17.0	52.0	31.0	83.0	57.0	87.0	13.0	0.00	13.0	6.50	[172]
Chinese (Singapore)	100	5.00	31.0	64.0	95.0	79.5	75.0	24.0	1.00	25.0	13.0	[172]
Chinese (Han)	111	9.00	35.1	55.9	91.0	73.4	73.8	24.3	1.80	26.1	14.0	[173]
Malay (ingapore)	100	2.00	22.0	76.0	98.0	87.0	79.0	20.0	1.00	21.0	11.0	[172]
Brazilian	143	55.9	35.7	8.40	44.1	26.2	74.1	23.8	2.10	25.9	14.0	[174]

[1]p<0,001

[2]p=0,05

Table 4. Genotype- and phenotype frequencies of the *SLCO1B1* intronic (T89595C) polymorphism in Roma, Hungarian and HapMap populations

Population	n	T89595C %				
		TT	TC	CC	TC+CC	C allele
Roma	470	308 (65.5)	150 (31.9)	12 (2.60)	162 (34.5)	0.185
Hungarian	442	285 (64.5)	141 (31.9)	16 (3.60)	157 (35.5)	0.196
European (CEU)	113	77 (68.1)	35 (31.0)	1 (0.90)	36 (31.9)	0.164
Italian	102	60 (58.8)	38 (37.3)	4 (3.90)	42 (41.2)	0.225
Indian (Gujarati, Houston)	101	88 (87.1)	13 (12.9)	0 (0.00)	13 (12.9)	0.064
Japanese (Tokio)	113	44 (38.9)	49 (43.4)	20 (17.7)	69 (61.1)	0.394
Chinese (Han)	135	44 (32.6)	60 (44.4)	31 (23.0)	91 (67.4)	0.452
Chinese (Colorado)	108	34 (31.5)	44 (40.7)	30 (27.8)	74 (68.5)	0.481
African (USA)	57	36 (63.2)	18 (31.6)	3 (5.30)	21 (36.9)	0.211
Kenya (Luhya)	109	78 (71.6)	29 (26.6)	2 (1.80)	31 (28.4)	0.151
Kenya (Maasai)	156	106 (67.9)	43 (27.6)	7 (4.50)	50 (32.1)	0.183
Nigerian (Yoruba)	147	109 (74.1)	36 (24.5)	2 (1.40)	38 (25.9)	0.136
Mexican (LA)	57	46 (80.7)	10 (17.5)	1 (1.80)	11 (19.3)	0.105

Table 5. Allele frequency of the *SLCO1B3* c.1683-5676A>G and c.334T>G polymorphisms in Roma, Hungarian and HapMap populations.

Population	1683-5676A>G		334T>G	
	A %	**G %**	**T %**	**G %**
Roma	**96.6**	**3.4**	**29.4**	**70.6**
Hungarian	**84.9**	**15.1**	**47.8**	**52.2**
European	85.3	14.7	14.3	85.7
Italian	89.7	10.3	11.4	88.6
Indian (Gujarati)	95.5	4.5	5.9	94.1
Mexican	89.7	10.3	12.9	87.1
African	82.5	17.5	51.8	48.2
African (Kenya)	75.0	25.0	69.5	30.5
African (Nigerian)	79.6	20.4	64.4	35.6
Chinese (Han)	81.4	18.6	26.6	73.4
Chinese (Colorado)	85.3	14.7	26.4	73.6
Japanese	84.5	15.5	29.9	70.1