New insights into genotype–phenotype correlation in individuals with different level of general non-specific reactivity of an organism

Alexander Mulik[1], Valery Novochadov[1], Alexander Bondarev[1], Sofya Lipnitskaya[1], Irina Ulesikova[1], Yulia Shatyr[1]

[1]Volgograd State University, Institute of Natural Sciences, Volgograd, Russia, http://www.volsu.ru/

Summary

The objective of the study was to investigate the genetic basis of general non-specific reactivity of an organism. Systematic search in PubMedCentral, PDB, KEGG and SNP databases identified a set of genes and their polymorphisms that can determine pain sensitivity and therefore the level of general non-specific reactivity of the human organism. Six SNPs were selected for genotyping kit design; 230 healthy volunteers were enrolled in the study. It was revealed that very high pain threshold was associated with allele A in rs1851048 and allele C in rs6777055. High level of general non-specific reactivity of an organism was associated with allele G in rs2562456 (OR=1.804, CI=1.139-2.857, p=0.011) and allele C in rs6923492 (OR=1.582, CI=1.071-2.335, p=0.021). Low level of general non-specific reactivity of an organism was associated with allele T in rs6923492 (OR=0.351, CI=0.154-0.799, p=0.010). A set of genes and single-nucleotide polymorphisms associated with the pain sensitivity and indirectly with the level of general non-specific reactivity of human organism were determined. The identified correlations reveal some molecular mechanisms of general non-specific reactivity of an organism variability and can guide further research in this area.

1 Introduction

The development of personalized medicine requires the search of integrative criteria reflecting the individual organization of human body homeostasis [1, 2]. The level of general non-specific reactivity of an organism (LGNRO) is an integrative characteristic, seen in the functional unity of all organism systems through the central coordination of their sensitivity, reactivity and activity [3]. It is well known that in the development of the phenotype variability the input of genetic factors is about 40%, 30% is determined by the previous ontogenesis and 30% is due to uniqueness of the combination of these factors at the time of their action [4, 5].

Previous studies revealed stable specific combinations of morphological and morpho-functional characteristics of the central nervous system, depending on individual LGNRO and variability of adaptive organism reserves [6]. Further research requires detailed study of the genetic basis of the LGNRO formation. As a universal indicator of LGNRO in experimental and human studies is pain threshold (PT), the task of identifying genes determining LGNRO can be reduced to the search for genes associated with pain sensitivity. Investigation of genetic basis of pain sensitivity and analgesia is very complex and urgent task [7]. The identification of PT genetic markers will provide understanding of molecular mechanisms, responsible for variation of organism reactivity, and can be used in personalized medicine.

This task can be achieved in two stages. The first step is to create a database of genes involved in the PT formation and reflecting the human LGNRO, and then select the most promising single nucleotide polymorphisms (SNP) in these genes by virtual screening and design an appropriate research kit. The second stage is a pilot study involving genotyping the most promising SNPs in the population of permanent residents of the Volgograd region, Russia.

2 Architecture/Implementation

A structured and easy-to-use database was created during bioinformatics research in the public domains: NCBI, PubMedCentral, PDB, KEGG and SNP databases. The following keywords were used: pain, pain threshold, pain sensitivity, and specific gene names in combination with: single nucleotide polymorphism, gene polymorphism. More than 400 sources in 10-year period were analyzed. The database was designed in Microsoft Access, database structure was visualized using Microsoft Visio (Fig. 1). The following criteria were used during the database creation: adequacy, comprehensiveness, stability. The database contains the following fields: the complete gene name, standard abbreviations and commonly used synonyms, a direct function of the encoded protein, the presence of polymorphisms, the functional effect of polymorphism (impact on the direct function of the encoded protein), described and possible effect on the reactivity, sources. SNP search was performed in Single Nucleotide Polymorphism Database [8].

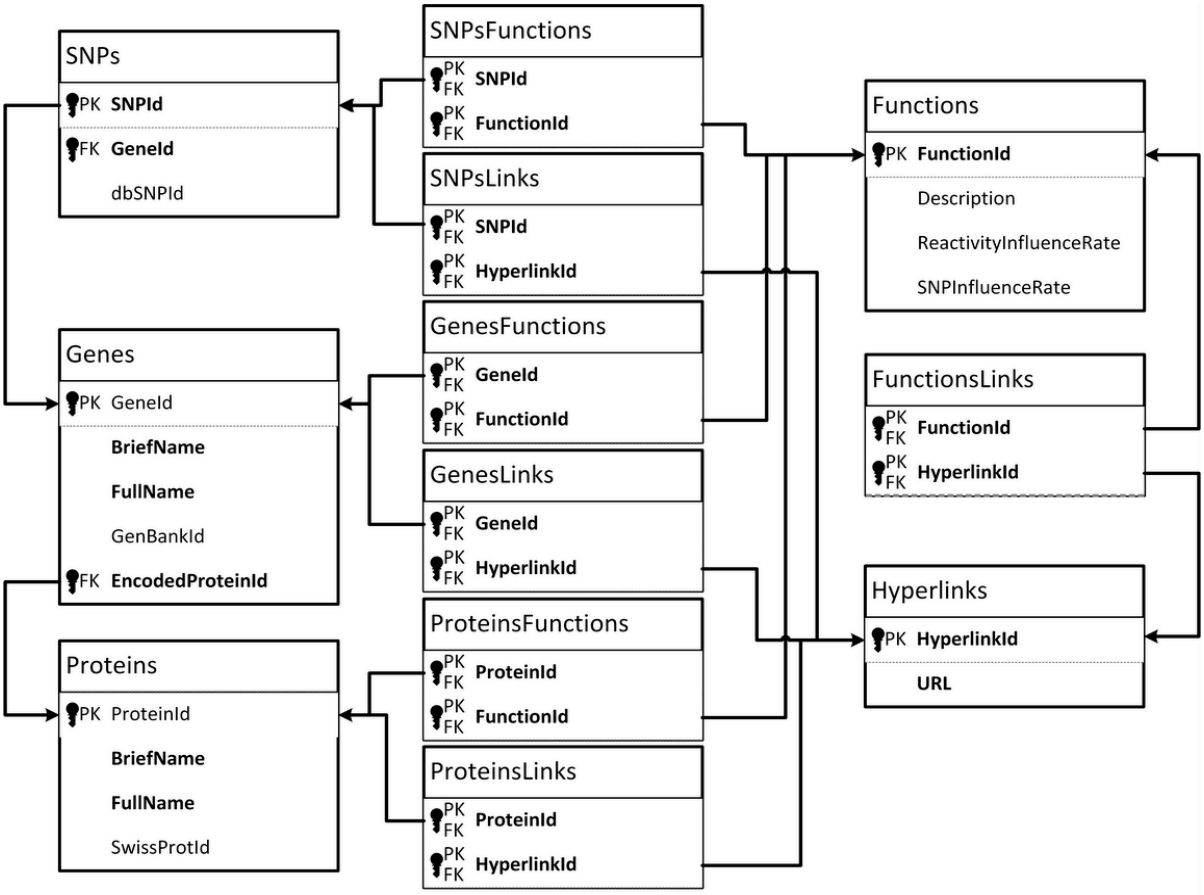

Figure 1: A block diagram of the database of genes associated with pain perception created in MS Vision

Two hundred and thirty clinically healthy men and women 18-23 years old were enrolled in the second part of the study (81 male and 149 female). They were selected by random sampling from 2250 full-time students of the Volgograd State University, Caucasians, and natives of the Volgograd region. The research was carried out in accordance with the "Universal Declaration on Bioethics and Human Rights", the informed consent for research and peripheral venous blood sampling was obtained.

LGNRO was evaluated for all participants by identifying PT by automatically measuring the time of a reflective hand removal from the light beam using device for determination of pain threshold "UgoBasile" (Italy) [3]. PT was registered in seconds at the time of the hand removal. High LGNRO corresponds to low PT values, low LGNRO – to high PT values, average LGNRO – to intermediate values of pain sensitivity.

Genotyping was performed by real-time PCR on Rotor Gene 6000 (Corbett Research, Australia). Primers and probes for this research were designed by Syntol (Russia). Genomic DNA was extracted from EDTA-stabilized peripheral venous blood by magnetic sorption method.

PT outliers were detected by the Grubbs test [9] used separately for men and women subgroups. The Shapiro-Wilk test for normality was used [10]. Comparison of genotype frequencies with predicted by Hardy-Weinberg law and with data from international databases was performed using the χ^2test with the confidence level 0.05.The association of genotypes with PT was studied using non-parametric Kruskal-Wallis test [11]. Correlation analysis was performed with Spearman rank correlation coefficient.

Calculations and graph plotting were performed in Statistica 8.0 (Statsoft Inc.) and MS Excel (v14.0) (Microsoft 2010) [12, 13].

3 Application

A block diagram of constructed database is shown in Figure 1. Completed database contains 164 human genes that met the following criteria: proven influence on the perception of pain, association with a variety of chronic pain syndromes (1), and the presence of annotated polymorphisms associated with these phenotypic variations (2). After exclusion of observations, which were not confirmed by further studies, cases of rare genetic syndromes or cases associated only with situations of severe visceral pain (cancer surgery, traumatic surgical effects, etc.) 24 genes were selected for further analysis (Table 1).

After NCBI-SNP database analysis 9 genes with non-synonymous SNPs with described phenotype variation were selected. These polymorphisms alter the amino acid sequence of the encoded protein which results in a change in protein conformation with a possible impact on its function. List of these polymorphisms is shown in Table 1 (second step).

Analysis of the influence of these SNPs on the secondary structure formation and stability and the probability of the protein functional properties alteration allowed us to select 6 promising SNPs: rs1851048 and rs6777055 in *cacna2d3* gene, encoding the voltage gated Ca2+ channels; rs2562456 in *znf-ld* gene of zinc-containing transcriptional regulator of DNA methylation; rs6923492 and rs362962 in *grm1*gene of metabotropic glutamate receptor; and rs6314 in *htr2a* gene, encoding serotonin receptor type 2A (Table 2).

Figure 2 shows a histogram of PT distribution. The distribution is asymmetrical with a large "tail" on the right and a trend to having multiple modes (Shapiro-Wilk test value 0.7508, $p<0.001$). It can be caused by the fact that some of the observations are significantly different from the main group.

Table 1: The list of genes and selected genetic polymorphisms associated with phenotypic characteristics of pain perception and reactive response to the pain perception in human

First step		Second step		
Gene	**Encoded protein**	**Gene**	**Records in NCBI-SNP**	**Most likely candidates**
The genes associated with the synthesis and reception of neurotransmitters				
adrb2	Adrenergic receptor β2			
comt	Catechol-O-methyltransferase	*comt*	3980	rs9265 rs74745580
gch1	GTP cyclohydrolase 1			
grm1	Glutamate receptor 1	*grm1*	74658	rs362829 rs6923492 rs362962
hcrtr2	Orexin B			
htr2a	Serotonin receptor 2A	*htr2a*	15514	rs6314 rs2274639
mc1r	Melanocortin-1 receptor			
oprd1	Opioid receptor, delta 1			
oprm1	Opioid receptor, mu 1	*oprm1*	28996	rs1799971 rs497976
p2rx7	Ca2+-permeable cationic channels			
slc6a4	Serotonin transporter	*slc6a4*	18642	rs28914822 rs25531
trpa1	Transient receptor potential A1			
trpv1	Transient receptor potential V1			
Genes associated with membrane transport of electrolytes				
cacna1a	Neuronal calcium channel			
cacna2d3	Voltage gated Ca2+ channels	*cacna2d3*	180000	rs6777055 rs1851048
kcns1	Voltage gated K+ channels			
scn9a	Voltage gated Na+ channels			
Genes associated with the synthesis of interleukins				
il10	Interleukin-10			
il1b	Interleukin-1β	*il1b*	2527	rs2853550 rs1799916
il6	Interleukin-6			
Tnfa	Tumor necrosis factor, α	*Tnfa*	802	rs267600955
Genes associated with certain metabolic response				
mt1x	Metallothionein			
vldlr	Very low density lipoprotein receptor			
znf-ld	Transcriptional regulator of DNA methylation	*znf-ld*	11	rs2562456

Table 2: Characteristics of selected SNPs

SNP	Alleles	MAF	Gene and its product
rs1851048	G/A	0.2232	*CACNA2D3, α2δ3*-subunit potential dependent calcium channel
rs6777055	A/C	0.1558	*CACNA2D3*, α2δ3-subunit of voltage-gated calcium channel
rs2562456	A/G	0.2198	*ZNF-LD*, is in linkage disequilibrium with the protein gene with domains "zinc finger" type 429
rs6923492	T/C	0.3960	*GRM1*, metabotropic glutamate receptor type I
rs362962	T/C	0.3898	*GRM1*, metabotropic glutamate receptor type I
rs6314	C/T	0.0747	*HTR2A*, serotonin receptor type 2A

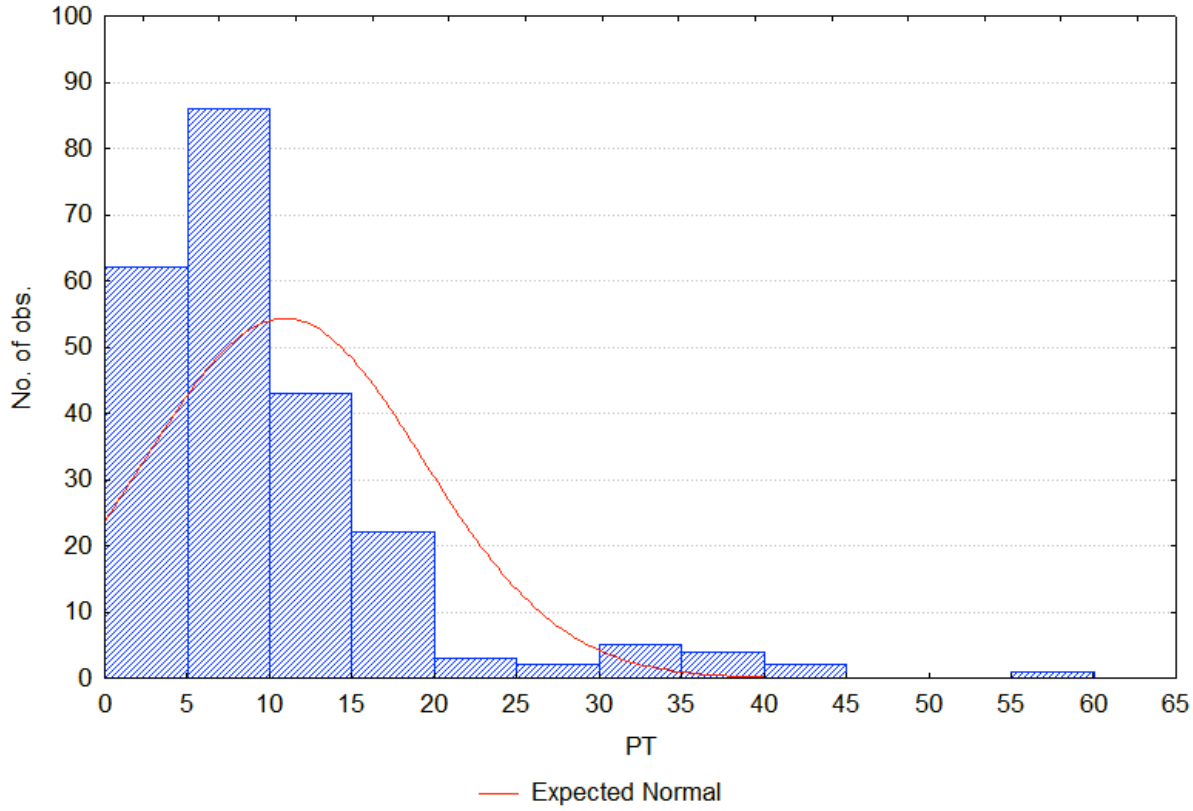

Figure 2: Distribution of PT values in the studied population before outlier removal

Using the Grubbs test separately for male and female subgroups revealed 15 outliers. The graph becomes unimodal, although doesn't fit normal distribution due to positive skewness and closeness of the lowest PT values to zero.

To assess the quality and representativeness of the data the compliance to Hardy-Weinberg equilibrium was checked. The observed allele and genotype frequencies are summarized in Table 3. The significance of differences between frequencies was checked by the $\chi 2$ test. For all studied polymorphisms genotype distribution frequency corresponded well to the one calculated by the Hardy-Weinberg equation.

When comparing the obtained allele frequencies with data of international genotyping projects (p2, Table 3) for three SNPs significant differences were revealed. This may indicate some differences in the frequency of allelic variants of these polymorphisms in the Volgograd region in comparison with other studied populations.

Table 3: Distribution of alleles and genotypes in the studied population.

SNP	Frequency of alleles		Frequency of genotypes			p1	p2
	Major	Minor	11	12	22		
rs1851048	0.6870	0.3130	0.4652	0.4435	0.0913	0.8946	**<0.0001**
rs6777055	0.8304	0.1696	0.6913	0.2783	0.0304	0.9837	0.4649
rs2562456	0.7391	0.2609	0.5478	0.3826	0.0696	0.9930	0.0581
rs6923492	0.5239	0.4761	0.2565	0.5348	0.2087	0.5508	**<0.0001**
rs362962	0.7804	0.2196	0.5913	0.3783	0.0304	0.2902	**<0.0001**
rs6314	0.9261	0.0739	0.8565	0.1391	0.0043	0.9699	0.953

p1 – $\chi 2$ test p-value for Hardy-Weinbergequilibrium check; p2 – $\chi 2$ test p-value for comparison the obtained allele frequencies with data of 1000 Genomes Project and the HapMap, p-values < 0.05 marked in bold.

Comparison of genotype frequencies in the main subgroup and in the outlier subgroup with abnormally high PT revealed significant association of polymorphisms rs1851048 and rs6777055 in *cacna2d3*gene with the probability of observing an abnormally high PT. The odds ratio for the presence of rs1851048 major allele was 0.236 (CI = 0.068-0.822, p = 0.0147), for rs6777055 major allele - 0.155 (CI =0.027-0.875, p = 0.0164). Thus the presence of minor alleles in these SNPs increases the probability of observing very high PT. The effect of rs6777955 on temperature sensitivity in Drosophila was described [14]. Thus, the structural features of the neuronal calcium channels may explain the observed relatively high PT in a small studied subgroup [15].

To study the effect of polymorphisms on PT we used nonparametric Kruskal-Wallis test. Table 4 shows the total p-values of the Kruskal-Wallis test, and p-values of individual comparisons between genotypes.

Table 4: Association with genetic markers of pain sensitivity threshold.

Marker	Kruskal-Wallis criterion			
	p-value	p 11-12	p 11-22	p 12-22
rs1851048	0.2717	1.0000	0.3937	0.3443
rs6777055	0.7146	1.0000	1.0000	1.0000
rs2562456	0.3466	1.0000	0.4419	0.6546
rs6923492	**0.0126**	1.0000	**0.0137**	**0.0449**
rs362962	0.2969	1.0000	0.4121	0.6590
rs6314	0.6460	1.0000	0.9484	0.9267
Sex	**0.0235**			

p-values< 0.05 marked in bold

There is a significant effect of rs6923492 polymorphism in the metabotropic glutamate receptor gene type I on the pain threshold. These differences are concentrated in subgroup homozygous for the minor allele, the major allele homozygotes and heterozygotes don't differ statistically. This fact demonstrates the complete dominance of the major allele of this polymorphism in effects on PT and LGNRO (Figure 3).

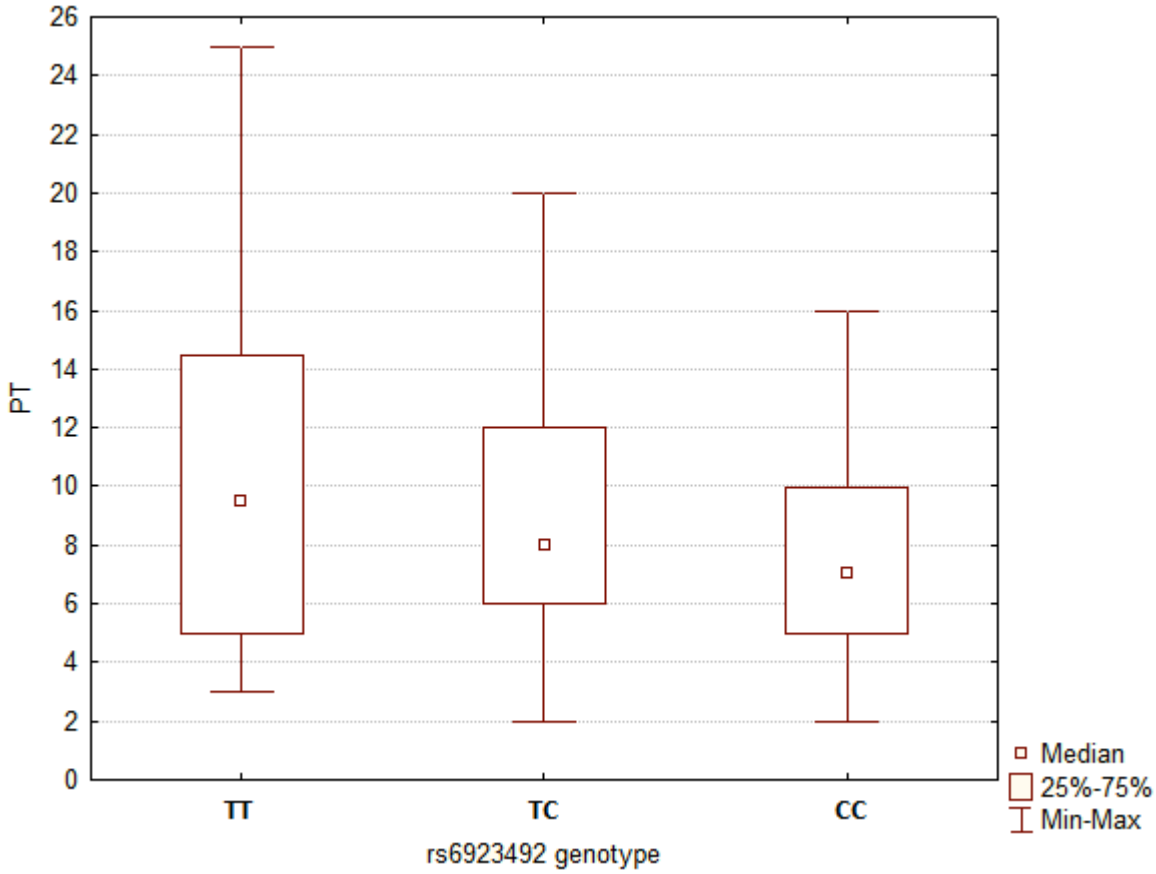

Figure 3: PT distribution for rs6923492 genotypes

According to current data rs6923492 is not associated with any disease state; this fact correlates well with its high frequency in the studied populations. So it is a good candidate for the role of one of the factors responsible for genetic variability of LGNRO in human.

Correlation between genotype, sex and PT was studied by the Spearman's rank correlation coefficient (Table 5).

Table 5: The matrix of Spearman correlation coefficient values for studied SNPs, sex and PT

	rs1851048	**rs6777055**	**rs2562456**	**rs6923492**	**rs362962**	**rs6314**	**Sex**	**PT**
rs1851048	1.0000	-0.0851	**0.1589**	-0.0208	-0.0039	0.0111	0.0058	-0.0497
rs6777055	-0.0851	1.0000	-0.0710	-0.1231	-0.0766	-0.0301	-0.0069	0.0433
rs2562456	**0.1589**	-0.0710	1.0000	0.0077	0.1011	0.0557	-0.0096	-0.0695
rs6923492	-0.0208	-0.1231	0.0077	1.0000	**0.2578**	0.0614	-0.0626	**-0.1862**
rs362962	-0.0039	-0.0766	0.1011	**0.2578**	1.0000	0.0436	-0.0282	-0.0767
rs6314	0.0111	-0.0301	0.0557	0.0614	0.0436	1.0000	0.0341	-0.0031
Sex	0.0058	-0.0069	-0.0096	-0.0626	-0.0282	0.0341	1.0000	**-0.1548**
PT	-0.0497	0.0433	-0.0695	**-0.1862**	-0.0767	-0.0031	**-0.1548**	1.0000

Coefficients with p-values < 0.05 marked in bold

Sex and rs6923492 show significant correlation with PT. The presence of the minor allele in this polymorphism of the metabotropic glutamate receptor gene type I is associated with a decrease in pain threshold and higher LGNRO. Higher PT was observed in men, which is consistent with previous studies.

Low values of correlation coefficients do not allow deriving equations for reliable predictions of LGNRO by the results of genotyping. This may be due to the integrative nature of LGNRO and its dependence on a large number of factors. Thus, the creation of such model requires the inclusion of additional genetic and / or phenotypic markers.

The highest correlation coefficient is found between rs6923492 and rs362962 in metabotropic glutamate receptor gene type I (R = 0.2538). This suggests that these polymorphisms are closely linked, and in the studied population certain combinations of these two mutations prevail. There is also a significant correlation between genotypes of rs1851048 and rs2562456, which can be explained by association of certain polymorphisms alleles in the most common haplotypes.

To convert the PT values into LGNRO the range was divided into three equal intervals. For high and low LGNRO the odds ratio for the presence of the variant allele was calculated. High LGNRO is associated with the presence of G allele of rs2562456 (OR=1.804, CI=1.139-2.857, p=0.011) and C allele of rs6923492 (OR=1.582, CI=1.071-2.335, p=0.021). Low LGNRO is associated with T allele of rs6923492 (OR=0,351, CI=0,154-0,799, p=0.010) (Table 6).

Table 6: Analysis of the frequency of occurrence of the studied SNPs with the level of reactivity

SNP	Gene	High LGNRO			Low LGNRO		
		OR	95% CI	p	OR	95% CI	p
rs1851048	*cacna2d3*	1.343	0.879-2.052	0.172	1.028	0.473-2.237	0.944
rs6777055	*cacna2d3*	0.953	0.567-1.604	0.858	1.204	0.477-3.044	0.693
rs2562456	*znf-ld*	**1.804**	**1.139-2.857**	**0.011**	0.504	0.189-1.341	0.163
rs6923492	*grm1*	**1.582**	**1.071-2.335**	**0.021**	**0.351**	**0.154-0.799**	**0.010**
rs362962	*grm1*	1.479	0.918-2.382	0.106	0.625	0.234-1.669	0.344
rs6314	*htr2a*	1.107	0.539-2.274	0.783	0.762	0.174-3.337	1.000

OR - odds ratio, 95% CI - 95% confidence interval, values with $p < 0.05$ marked in bold

Thus usage of LGNRO instead of primary PT data confirms the role of rs6923492, and reveals additional association with *ZNF-LD* gene.

It should be noted that, statistically significant results were obtained for SNPs with the highest frequency of the minor allele, as this provides the maximum power of statistical tests. Our data can't rule out the association of PT and LGNRO with other studied SNPs because of the small number of subgroups. This will require larger studies.

4 Discussion

The metabotropic glutamate receptor is directly related to reactivity because the glutamatergic system activation provides trophic supply of the organism response to external factors. Therefore high LGNRO must comply with relatively more intensive work of the glutamatergic system. Both studied *grm1* polymorphisms are localized in exon 8 of the gene encoding the topological receptor region which is located in the neuron cytoplasm (XP_011534086.1:p.Ser993Pro). Due to the fact that the binding sites for glutamate are closer to the N-end of the molecule, it is better annotated and three dimensional configuration in PDB is present only for the first 860 amino acids. This fact complicates the search for the mechanism of this molecule function changes due to studied SNPs. Rs6923492 substitutes serine (dominant C allele) by proline (recessive T allele) in position 993. Beginning with 988

position the GRM1 molecule has a proline-rich site: it occurs 11 times out of 37 next amino acids. It is believed that replacement Ser/Pro at the beginning of a proline-rich site further changes its conformation and, according to the phenotypic effect, decreases the receptor affinity to the G protein-coupled receptor kinase and, indirectly, increases intracellular response to the reception of glutamate [16, 17].

This effect should be interpreted as an increase in reactivity due to the central glutamatergic mechanisms of its regulation. Rs362962 is located in intron, so its effect will probably relate to regulation of *grm1*gene expression and so far is unclear.

Polymorphism rs2562456 associated with high LGNRO is found on ZNF429 gene site, located on chromosome 19 in the area previously described as an unknown gene LOC400680. There are more than 20 SNPs in this region, coding a protein of 674 amino acids, which corresponds to DNA-binding transcriptional regulators with "zinc finger" motif. The three dimensional structure of this protein is not deciphered yet [18].We can only hypothesize that the presence of the G allele changes the overall ability of ZNF429 to activate transcription because reactivity can be realized through the synthesis of specific molecules. It is confirmed by the findings in which ZNF429 structure variability was engaged in such phenotypic properties of the organism as pain sensitivity and the overall tolerability to pain [19, 20].

Despite the complexity of the interpretation of found SNPs effects on PT and LGNRO, these data show the feasibility of used algorithm: virtual gene screening, virtual SNP screening, design and synthesis of genotyping kit, revealing genotype-phenotype correlations, interpretation of molecular and physiological mechanisms of the effects.

Acknowledgements

The study was carried out with financial support of Russian Foundation for Basic Research in the framework of a research project number 15-06-08034.

References

[1] B. Paul, S. Barnes, W. Demark-Wahnefried, et al. Influences of diet and the gut microbiome on epigenetic modulation in cancer and other diseases. *Clinical Epigenetics*, 7:e112, 2015

[2] K. Suratanee, D.D.A. Plaimas: A novel network-based scoring method to identify disease–disease associations. *Bioinformatics and Biology Insights*, 9:175-186, 2015

[3] A.B. Mulik. General Non-Specific Reactivity of an Organism: development, estimation, application. Volgograd, VolSU, 2001. – 144 p.

[4] D.M. Gash, A.S. Deane Neuron-based heredity and human evolution. Frontiers in Neuroscience, 9:e209, 2015; L. Marmiesse, R. Peyraud, L. CottretFlexFlux: combining metabolic flux and regulatory network analyses. *BMC Systems Biology*, 9:e93, 2015

[5] J. Weeland, G. Overbeek, B.O. de Castro, W. Matthys Underlying mechanisms of gene–environment interactions in externalizing behavior: a systematic review and search for theoretical mechanisms. *Clinical Child and Family Psychology Review*, 18(4):413-442, 2015

[6] A.B. Mulik, U.A. Shatyr, M.V. Postnova, N.O. Nazarov. Adaptive capacity of individuals with different the general non-specific reactivity level under standard environmental load. *Bulletin of Volgograd State University*. Series 11: Natural sciences, 2:91-96, 2012

[7] M.L. LaCroix-Fralish, J.S. Mogil. Progress in Genetic Studies of Pain and Analgesia. *Annu. Rev. PharmacolToxicol*. 2009 ; 49: 97–121. doi:10.1146/annurev-pharmtox-061008-103222.

[8] S.T. Sherry, M. Ward, K. Sirotkin. dbSNP - database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Research*, 9(8): 677–679, 1999

[9] F.E. Grubbs. Sample criteria for testing outlying observations. *Annals of Mathematical Statistics*, 21(1):27–58, 1950

[10] S.S. Shapiro, M.B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52 (3–4): 591–611, 1965

[11] W.H. Kruskal, A. Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260):583–621, 1952

[12] Afifi, S. May, V.A. Clark. Computer-Aided Multivariate Analysis. *CRC Press: Chapman & Hall/CRC Texts in Statistical Science*, 2003

[13] R. de Levie. Advanced Excel for Scientific Data Analysis. Science: *Oxford University Press*, 2004

[14] G.G. Neely, A. Hess, M. Costigan. A genome-wide Drosophila screen for heat nociception identifies alpha2delta3 as an evolutionarily conserved pain gene. *Cell*, 143(4): 628–638, 2010

[15] Peyrache, M.M. Lacroix, P. Petersen, G. Buzsáki. Internally-organized mechanisms of the head direction sense. *Nat. Neurosci*, 18(4):569–575, 2015

[16] A.C. Emery, S., Pshenichkin, G.R. Takoudjou, E. Grajkowska, B.B. Wolfe. The protective signaling of metabotropic glutamate receptor 1 is mediated by sustained, beta-arrestin-1-dependent ERK phosphorylation. *Journal of Biological Chemistry*, 285: 26041–26048, 2010

[17] M.S. Mehta, S.C.Dolfi, R. Bronfenbrener, E. Bilal., C.Chen, D. Moore. Metabotropic glutamate receptor 1 expression and its polymorphic variants associate with breast cancer phenotypes. *PLoS One*, 8(7):e69851, 2013

[18] H. Kim, E. Ramsay, H. Lee. Genome-wide association study of acute post-surgical pain in humans. *Pharmacogenomics*, 10(2):171–179, 2009

[19] E.E. Young, W.R. Lariviere, I.Belfer. Genetic basis of pain variability: recent advances. *Journal of Medical Genetics*. 49(1):1-9, 2012

[20] M.A. Petrie, M. Suneja, E. Faidley, R.K. Shields. A minimal dose of electrically induced muscle activity regulates distinct gene signalling pathways in humans with spinal cord injury. *PLoS One*, 9(12):e115791, 2014