Meta Gene 2 (2014) 283–298

Contents lists available at ScienceDirect

Meta Gene

PSMA6 (rs2277460, rs1048990), *PSMC6* (rs2295826, rs2295827) and *PSMA3* (rs2348071) genetic diversity in Latvians, Lithuanians and Taiwanese

Tatjana Sjakste [a,*], Natalia Paramonova [a], Lawrence Shi-Shin Wu [b], Zivile Zemeckiene [c], Brigita Sitkauskiene [d], Raimundas Sakalauskas [d], Jiu-Yao Wang [e], Nikolajs Sjakste [f,g]

[a] *Genomics and Bioinformatics, Institute of Biology of the University of Latvia, Miera str. 3, LV2169, Salaspils, Latvia*
[b] *Institute of Biomedicine, Tzu-Chi University, Tzu-Chi, Taiwan*
[c] *Department of Laboratory Medicine, Medical Academy, Lithuanian University of Health Sciences, Kaunas, Lithuania*
[d] *Department of Pulmonology and Immunology, Medical Academy, Lithuanian University of Health Sciences, Kaunas, Lithuania*
[e] *Division of Allergy and Clinical Immunology, College of Medicine, National Cheng Kung University, Tainan, Taiwan*
[f] *Faculty of Medicine, University of Latvia, Riga, Latvia*
[g] *Latvian Institute of Organic Synthesis, Riga, Latvia*

ARTICLE INFO

Article history:
Received 26 February 2014
Revised 11 March 2014
Accepted 17 March 2014
Available online 17 April 2014

Keywords:
PSMA3
PSMA6
PSMC6
Proteasome
SNP
Genetic diversity
Human population

ABSTRACT

PSMA6 (rs2277460, rs1048990), *PSMC6* (rs2295826, rs2295827) and *PSMA3* (rs2348071) genetic diversity was investigated in 1438 unrelated subjects from Latvia, Lithuania and Taiwan. In general, polymorphism of each individual locus showed tendencies similar to determined previously in HapMap populations. Main differences concern Taiwanese and include presence of rs2277460 rare allele A not found before in Asians and absence of rs2295827 rare alleles homozygotes TT observed in all other human populations. Observed patterns of SNPs and haplotype diversity were compatible with expectation of neutral model of evolution. Linkage disequilibrium between the rs2295826 and rs2295827 was detected to be complete in Latvians and Lithuanians ($D' = 1$; $r^2 = 1$) and slightly disrupted in Taiwanese ($D' = 0.978$; $r^2 = 0.901$).

Abbreviations: LV, Latvian population; LT, Lithuanian population; TW, Taiwanese population; UPS, ubiquitin–proteasome system; SNP, single nucleotide polymorphism; TF, transcription factor; TFBS, transcription factor binding site; T2DM, type 2 *diabetes mellitus*; HapMap-CEU, NorthWestern Europeans; HapMap HCB, Han Chinese; HapMap JPT, Japanese; HWE, Hardy–Weinberg equilibrium; LD, linkage disequilibrium.

* Corresponding author. Tel.: +371 29716053; fax: +371 67944988.
E-mail address: tanja@email.lubi.edu.lv (T. Sjakste).

http://dx.doi.org/10.1016/j.mgene.2014.03.002

Population differentiation (F_{ST} statistics) was estimated from pairwise population comparisons of loci variability, five locus haplotypes and *PSMA6* and *PSMC6* two locus haplotypes. Latvians were significantly different from all Asians at each of 5 SNPs and from Lithuanians at the rs1048990 and *PSMC6* loci. Lithuanian and Asian populations exhibited similarities at the *PSMC6* loci and were different at the *PSMA6* and *PSMA3* SNPs. Considering five locus haplotypes all European populations were significantly different from Asian; Lithuanian population was different from both Latvian and CEU.

Allele specific patterns of transcription factor binding sites and splicing signals were predicted *in silico* and addressed to eventual functionality of nucleotide substitutions and their potential to be involved in human genome evolution and geographical adaptation. Current study represents a novel step toward a systematic analysis of the proteasomal gene genetic diversity in human populations.

Introduction

The ubiquitin-proteasome system (UPS) is the major nonlysosomal proteolytic pathway affecting crucial intracellular processes. UPS deregulation has been implicated in the efficiency of the immune response, ageing, inflammatory and many pathological processes (Sorokin et al., 2009; Willis et al., 2010; Zemeckienė et al., 2013).

UPS components possess potential to be a therapeutic target for treatment of several diseases (Bedford et al., 2011).

Importance of proteasomes in both normal and pathological processes triggers interest for search of sequence variations in the proteasomal genes to be associated with human pathologies including cardiovascular disorders (Alsmadi et al., 2009; Banerjee et al., 2008; Barbieri et al., 2008; Bennett et al., 2008; Ozaki et al., 2006; Sjakste et al., 2007b), *diabetes mellitus* (Sjakste et al., 2007a, 2007b), autoimmune diseases (Sjakste et al., 2004, 2010, in press; Trapina et al., 2009), children obesity (Kupca et al., 2013), cancer and its outcome (Bachmann et al., 2010). However, these associations could significantly vary in different ethnic populations as it was shown for coronary artery disease (Wang et al., 2013).

No systematic analysis has been done until now to evaluate the proteasomal gene genetic diversity in humans on population level. It appears that genotyping of the *PSMB5* gene single nucleotide polymorphism (SNP) in four American ethnic groups (Wang et al., 2008) and the 14q13.2 microsatellite polymorphism in Latvian and Finland populations (Kalis et al., 2002; Sjakste et al., 2007a) are the only studies directly addressed to that objective. Case/control disease association studies could provide significant information on population diversity. The rs1048990 of the *PSMA6* gene was widely genotyped during the last decade in several European and Asian populations for association with cardiovascular disorders, type 2 diabetes mellitus and other pathologies and their outcome (Table 1 and Wang et al. (2013) for references). Several SNPs located within and upstream the *PSMA6* gene as well as within the *PSME1*, *PSME2* and *PSMA3* genes were genotyped in Latvians on susceptibility to different pathologies (Kupca et al., 2013; Sjakste et al., 2007b, in press; Trapina et al., 2009). HapMap and PERLEGEN projects (http://www.ncbi.nlm.nih.gov/snp/) provide significant information on the genetic diversity of many individual loci in different ethnic populations; however this information is based on analysis of very small subject groups, more detailed studies are necessary.

In the current study five SNPs belonging to the *PSMA6* (rs2277460 and rs1048990), *PSMC6* (rs2295826 and rs2295827) and *PSMA3* (rs2348071) genes, have been genotyped on genetic diversity in Latvian (LV), Lithuanian (LT) and Taiwanese (TW) populations to evaluate extent of diversity and population differentiation. Allele specific patterns of transcription factor binding sites (TFBSs) and splicing signals were predicted *in silico* to reveal a potential of nucleotide substitutions in proteasomal genes to be important for human genome evolution and adaptation.

Table 1
Polymorphism description.

Marker ID	Coordinates cytogenetic/ genomic	Gene	Function	MAF/MA count	Association findings	
					Disease	Reference
rs2277460	14q13.2:/14:35761573	*PSMA6*	Promoter: c.-110C>A (c.-109-1C>A)	A = 0.0634/138	T2DM JIA	Sjakste et al., 2007b Trapina et al., 2009, Sjakste et al., in press
rs1048990	14q13.2/14:35761675	*PSMA6*	5′-UTR: c.-8C>G	G = 0.195/425	T2DM CVD	Sjakste et al., 2007b; Barbieri et al., 2008 Ozaki et al., 2006; Sjakste et al., 2007a, 2007b; Takashima et al., 2007; Bennett et al., 2008; Liu et al., 2009; Banerjee et al., 2009; Alsmadi et al., 2009; Hinohara et al., 2009; Honcharov et al., 2009; Ikeda et al., 2012; Heckman et al., 2013; Wang et al., 2013
					Cancer	Bachmann et al., 2010
rs2295826	14q22.1/14:53174923	*PSMC6*	Intron 1: c.128-104A>G	G = 0.124/271	JIA	Sjakste et al., in press
rs2295827	14q22.1/14:53174981	*PSMC6*	Intron 1: c.128-46C>T	T = 0.102/222	JIA	Sjakste et al., in press
rs2348071	14q23/14: 58730626	*PSMA3*	Intron 7: c.543 + 138G>A; c.522 + 138G>A; c.576 + 138G>A	A = 0.369/804	JIA Obesity	Sjakste et al., in press Kupca et al., 2013

Ancestral allele is underlined in the motif description. MAF/MA count is given according to the 1000 genome phase 1 population project data Abbreviations: MAF — minor allele frequency; MA — allele minor in Caucasians; T2DM — type 2 diabetes mellitus; CVD — cardiovascular disorders; JIA — juvenile idiopathic arthritis.

Materials and methods

Samples

LV population was represented by 191 (117 females) patients of Riga Bikernieki Hospital, specialized in trauma medicine. Patients with only trauma diagnosis and without autoimmune and cardiovascular disorders, type 2 *diabetes mellitus* (T2DM) and obesity were included in this cohort. LT population was represented by 150 individuals (97 females) underwent prophylactic evaluation at Kaunas Family Medicine Centres and Hospital of Lithuanian University of Health Sciences and being without diagnosis or familial predisposition to congenital diseases, acute or chronic infections, oncological, autoimmune or any other chronic diseases, immunodeficiency and obesity. TW study subjects enrolled from elementary school for allergy diseases screen were represented by 1097 children (558 girls) without asthma and asthma history. Written informed consent was obtained from all the participants. LV, LT and TW studies were approved by the Central Medical Ethics Commission of the Latvian Ministry of Health, Kaunas Regional Biomedical Research Ethics Committee and Ethics Committee of Taoyuan General Hospital respectively.

DNA extraction and genotyping

Genomic DNA was extracted from nucleated blood cells using a kit for genomic DNA extraction (Fermentas, Vilnius, Lithuania) or QIAamp DNA blood mini kit (Qiagen, Germany) and from oral swabs using QIAamp DNA Mini Kit (QIAGEN, Valencia, CA, USA) according to the manufacturer's protocols. Quality and quantity of DNA were determined using agarose gel electrophoresis and spectrophotometry. DNA samples were stored at −80 °C until use.

Table 1 summarizes information on the polymorphic loci genotyped in our study.

Genotyping methods and primer sequences used by LV and LT teams are indicated in Supplementary Table 1. In brief, basic PCR was performed using the DreamTaq polymerase (Fermentas, Vilnius, Lithuania) with following parameters: 94 °C for 5 min; then 35–40 cycles of 94 °C for 45 s, appropriate annealing temperature (55–61 °C) for 45 s, 72 °C for 45 s and a final extension at step at 72 °C for 7 min. DNA digestion by restriction enzymes was performed according to the producer's protocols (Fermentas, Vilnius, Lithuania). Allele specific amplification used to identify the rs2277460 alleles was followed by the *RsaI* digestion to genotype the rs1048990. The rs2295826 and rs2295827 were typed simultaneously in one *DdeI* digestion reaction. Amplified and digested products were analysed by electrophoresis in 1–3% agarose gel. TW team genotyped SNPs using high throughput, 384-microtiter plate, MassARRAY™ System, SEQUENOM according to the manufacturer's instructions. In brief, DNA containing the SNP site of interest was amplified, followed by the homogenous MassEXTEND™ assay in which label-free primer extension chemistry was used to generate allele-specific diagnostic products of unique molecular weight suitable to be distinguished through the application of matrix assisted laser desorption ionization time-of-flight mass spectrometry. For quality control 16 randomly chosen samples per each marker were genotyped in duplicate in different experiments in each LV, LT and TW population. The concordance of the genotyping was 100%. Genotyping data were verified by direct sequencing of the corresponding DNA fragments in both directions using the Applied Biosystems 3130xl Genetic Analyzer.

Alleles and genotype frequencies for the rs2277460 (ss24557113), rs1048990 (ss35076445), rs2295826 (ss3239727 and ss69157456), rs2295827 (ss23619651) and rs2348071 (ss3302481) were extracted from public available dbSNP (build 13) entries at NCBI (http://www.ncbi.nlm.nih.gov/snp) for 47, 40 and 43 unrelated participants of HapMap-CEU (NorthWestern European), HCB (Han Chinese) and JPT (Japanese) populations respectively and 5 locus genotypes were reconstructed for each individual.

Loci description and nucleotide numbering are given according to the recommended nomenclature system (http://www.genomic.unimelb.edu.au/mdi/mutnomen/recs.html/). The chromosome 14 GRCh37.p5 assembly (NCBI reference sequence: NC_000014.8) sequence information was used for loci description, nucleotide numbering and primer design that was done using the Primer 3.0 program.

Data analysis

Personalised genotyping data documentation resulted in knowledge of 5 locus genotype (5-LG: rs2277460/rs1048990/rs2295826/rs2295827/rs2348071) of each individual participant of the study. The 5-LGs, observed

haplotypes, genotypes of each individual locus (single locus genotype or SLG) and alleles' frequencies were estimated by direct gene counting and deviations from Hardy–Weinberg equilibrium (HWE) were tested by χ^2 test (Rodriguez et al., 2009; http://www.oege.org/software/hwe-mr-calc.shtml). DnaSP[21] version 5 (Rozas, 2009; http://www.ub.es/dnasp/) was used to reconstruct the haplotypes from un-phased genotypes, evaluates the nucleotide diversity, performs Tajima's D (Tajima, 1989) and Fu and Li's F^* and D^* (Fu and Li, 1993) tests of neutrality, and evaluates the pairwise linkage disequilibrium (LD) between the loci (D and r^2) and pairwise population differentiation (F_{st}). Haplotype age and phylogenetic relationships were obtained using the Reduced Median algorithm of the Phylogenetic network software Fluxus 4.611 (http://www.fluxus-engineering.com).

SNPs functional analysis in silico

An eventual functional significance of nucleotide substitutions was analysed *in silico* on the sequence similarity to TFBSs using Genomatix software, MatInspector, release 7.4 online tool (Cartharius et al., 2005; www.genomatix.de/) with core and matrix similarity of 1.000 and more than 0.800 respectively. Splicing signals were predicted by Human Splicing Finder version 2.4 (Desmet et al., 2009; http://www.umd.be/HSF/) with standard threshold values for branch point, donor and acceptor splice sites, enhancer, silencer, hnRNP and other splicing motifs.

Results

SNPs' diversity

Data on SNPs allele and genotype presentation in LV, LT and TW populations are presented in Table 2. In general, alleles' and genotypes' distribution was similar between LV, LT and CEU and TW, CHB and JPT populations respectively (data for HapMap populations are not shown). However, the rs2277460 rare allele A being previously not found in Asians, was observed in Taiwanese; and the rs2295827 rare allele TT homozygotes found in all human populations studied previously, were not detected in Taiwanese at all. The rs2348071 allele G was major in both LV and LT populations similar to CEU and allele A was major allele in TW similar to HCB and JPT. The rs1048990 minor allele was observed more than twice frequently in TW than in each LV and LT ($P < 0.001$) and significantly more frequent in LT compared to LV ($P < 0.05$). Minor alleles at the rs2295826 and rs2295827 were significantly ($P < 0.05$) less frequent in LV than in each of LT and TW and alleles' and genotypes' presentation was found to be similar between LT and TW ($P > 0.05$). TW was found to deviate significantly

Table 2
SNPs diversity in Latvian, Lithuanian and Taiwanese populations.

Marker ID	Population	MAF	Genotype frequency			HWP (χ^2)	Nucleotide diversity (π)	Test of neutrality		
			11	12	22			Tajima's D_T	Fu and Li's D^*	Fu and Li's F^*
rs2277460	LV	0.0654	0.8691	0.1309	–	0.94	0.12265	−0.18389	0.42624	0.27631
	LT	0.0933	0.8133	0.1867	–	1.59	0.16981	0.06301	0.43609	0.37690
	TW	0.0073	0.9854	0.0146	–	0.06	0.01449	−0.66360	0.37096	0.03623
rs1048990	LV	0.0890	0.8272	0.1676	0.0052	0.21	0.16259	0.05540	0.42624	0.36631
	LT	0.1533	0.7067	0.2800	0.0133	0.92	0.26051	0.60431	0.43609	0.57821
	TW	0.3254	0.4211	0.5068	0.0702	26.15***	0.43925	1.98494	0.37096	1.09877
rs2295826	LV	0.1047	0.8115	0.1673	0.0209	2.16	0.18799	0.20752	0.42624	0.42353
	LT	0.1700	0.7000	0.2600	0.0400	0.93	0.28314	0.73936	0.43609	0.62844
	TW	0.1545	0.7092	0.2726	0.0182	2.05	0.26140	0.87595	0.37096	0.65397
rs2295827	LV	0.1047	0.8115	0.1673	0.0209	2.16	0.18799	0.20752	0.42624	0.42353
	LT	0.1700	0.7000	0.2600	0.0400	0.93	0.28314	0.73936	0.43609	0.62844
	TW	0.1468	0.7065	0.2935	–	34.45***	0.25056	0.80841	0.37096	0.62677
rs2348071	LV	0.2932	0.5340	0.3456	0.1204	5.28*	0,41555	1.57068	0.42624	0.93621
	LT	0.2667	0.5400	0.3867	0.0733	0.02	0.39242	1.39150	0.43609	0.87098
	TW	0.6541	0.1859	0.3199	0.4941	94.14***	0.45274	2.06904[nd]	0.37096	1.13251

LV, LT and TW abbreviations are used to indicate Latvian, Lithuanian and Taiwanese populations respectively; MAF is given according to allele being minor in European populations. Statistical significance levels of probability correspond to: [nd] not determined ($P = 0.05$); * $P < 0.05$; *** $P < 0.001$.

($P < 0.001$) from HWE at the rs1048990 and rs2295827 where rare allele homozygotes appear to be underrepresented and at the rs2348071 where heterozygotes appear to be underrepresented. Slight underrepresentation of the rs2348071 heterozygotes ($P < 0.05$) was also detected in LV.

To assess whether the observed patterns of SNPs diversity corresponded to expectations under the neutral model of evolution, the D_T, D^* and F^* statistics were applied and data are presented in Table 3. For LV and TW, D_T was negative for the rs2277460; in all other cases D_T was positive. At the rs2348071, D_T was slightly more than 2 in TW, however, it did not reach level of significance ($P = 0.05$). For all populations at all loci, D^* and F^* were positive and no D_T, D^* and F^* values were significant.

Linkage disequilibrium

Pairwise LD was estimated by DnaSP software and evaluated using D and r^2 parameters (Table 3). The rs2295826 and rs2295827 were found to be in complete LD ($D = 1$, $r^2 = 1$) in LV and LT; linkage was slightly disrupted in TW ($D = 0.978$, $r^2 = 0.901$).

Haplotype diversity

Spectra and frequencies of five locus haplotypes are given in Table 3 for LV, LT, TW and HapMap CEU, HCB and JPT populations and listed in order of frequency decrease in LV. From 18 haplotypes inferred by DnaSP, the first 10 listed showed frequency higher than 5% in one population at least. Haplotypes Hap11–12 and Hap13–18 were inferred only for LT and TW respectively and appear to be rare (≤2%) in human populations over the world.

Table 3
Allelic composition, frequency and other characters of the 5 locus haplotypes in human populations.

Haplotype		Populations (2n)					
ID	1–2–3–4–5	LV (382)	LT (300)	CEU (94)	TW (2194)	HCB (80)	JPT (86)
Hap 1	CCACG	0.5576	0.4967	0.6064	0.2388	0.1875	0.2558
Hap 2	CCACA	0.2068	0.1467	0.2021	0.3464	0.3000	0.2209
Hap 3	CCGTG	0.0681	0.0833	0.0638	0.0264		
Hap 4	ACACA	0.0393	0.0267	0.0106	0.0027		
Hap 5	CGACA	0.0340	0.0067	0.0745	0.1796	0.2625	0.2791
Hap 6	CGACG	0.0314	0.1100	0.0426	0.0706	0.0625	0.0581
Hap 7	ACACG	0.0262	0.0400		0.0046		
Hap 8	CGGTG	0.0236				0.0625	0.0116
Hap 9	CCGTA	0.0131	0.0300		0.0510	0.1250	0.1163
Hap 10	CGGTA		0.0333		0.0643		0.0581
Hap 11	ACGTA		0.0233				
Hap 12	AGACG		0.0033				
Hap 13	CGGCA				0.0068		
Hap 14	CCGCG				0.0023		
Hap 15	CGGTG				0.0023		
Hap 16	CGATA				0.0018		
Hap 17	CCGCA				0.0014		
Hap 18	CCATG				0.0009		
Hd		0.638	0.710	0.586	0.779	0.792	0.797
Test of neutrality							
Tajima's D, D_T		0.71472	1.35443	−0.20222	2.00713nd	1.85362$^{\#*}$	1.91830$^{\#*}$
Fu and Li's D^*		0.93163	0.94926	−0.01266	0.82277	0.96034	0.95389
Fu and Li's F^*		1.02654	1.29964	−0.08777	1.52548$^{\#*}$	1.45955	1.1.48121$^{\#*}$
LD 3–4							
D′		1.000	1.000	1.000	0.978	1.000	1.000
r^2		1.000***B	1.000***B	1.000***B	0.901***B	1.000***B	1.000***B

The loci 1, 2, 3, 4, and 5 in haplotypes' configurations correspond in sequence to the rs2277460, rs1048990, rs2295826, rs2295827 and rs2348071 loci respectively; Hd — haplotype diversity; LD — linkage disequilibrium. Statistical significance levels of probability correspond to: $^{\#*}$ $0.10 > P > 0.05$; nd not determined: $P = 0.05$; *** $P < 0.001$; B — Bonferroni correction for multiple tests.

Patterns of haplotypes distribution are drastically different between Europeans and Asians. Hap1 being the most frequent haplotype in Europeans is only second in TW and JPT and third in HCB; Hap5 being minor in Europeans is the first in JPT, second in HCB and third in TW. More examples could be listed (see Table 3). Pattern of diversity being similar between LV and CEU is different to some extent from LT. The most impressive difference concerns the Hap6 being minor (less than 5%) in both LV and CEU and presented in 11% of Lithuanians. Level of haplotype diversity is approximately the same in TW, HCB and JPT (0.779, 0.792, and 0.797 respectively); however minor in TW Hap3, Hap4, Hap7 and Hap13–18 were not suggested for HCB and JPT; Hap8 predicted for HCB and JPT was not inferred for TW. It is of interest that Hap10 absent in LV, CEU and HCB, was observed in LT, TW and JPT with similar frequency (3–6%).

Haplotypes' distributions were analysed on neutrality of evolution by D_T, D^* and F^* statistics (Table 3). D_T slightly exceeded 2, was obtained only for TW, however statistical significance was not determined ($P = 0.05$); D_T values for other populations as well as D^* and F^* values for all populations were not statistically significant. The *PSMA6* (rs2277460/rs1048990) and *PSMC6* (rs2295826/rs2295827), two locus haplotype patterns were also tested on neutrality and no D_T, D^* and F^* values were significant for all populations (data not shown).

Fig. 1 illustrates proportions and phylogenetic relationships between haplotypes being most frequent (frequency more than 1%) in particular population. Hap2 having ancestor alleles at all five loci, was considered to be a common ancestor in all populations analysed. That haplotype being one of most wide-spread in modern Asians appears to be an immediate ancestor of the Hap1, Hap4, Hap5 and Hap9. Substitution A to G at the rs2348071 appears to be the most ancient among the analysed mutations. That mutation appears to happen about 8000–15,000 years ago and generated the Hap1, the predominant haplotype in modern Europeans. Major in Europeans and one of the most frequent in Asians, the Hap1 appears to be the immediate ancestor for Hap3, Hap6 and Hap7 having more recent time of origin.

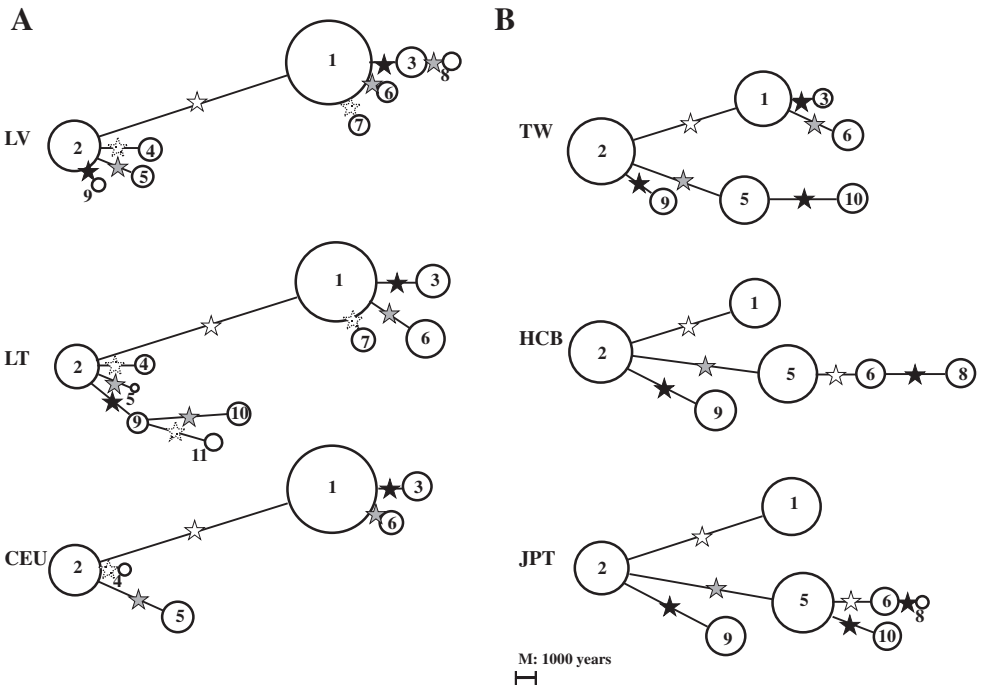

Fig. 1. Median joining network of *PSMA6*/*PSMC6*/*PSMA3* haplotypes. The network was generated using the Reduced Median algorithm of the Phylogenetic network software Fluxus 4.611 (http://www.fluxus-engineering.com). Panel A: LV — Latvians, LT — Lithuanians, CEU — NorthWestern Europeans; Panel B: TW — Taiwanese, HCB — Han Chinese, JPT — Japanese. Haplotypes are numbered as in Table 3. Mutations at the rs2277460, rs1048990, rs2295826&rs2295827 and rs2348071 are indicated as stars of interrupted outline, grey, black and white colours respectively. Length of the branches is equal to time of generation and dissemination of progeny haplotype.

Population differentiation

Patterns of population differentiation were analysed using F_{ST} statistics for SNPs' and haplotype frequencies and results are presented in Table 4. Locus-by-locus pairwise comparison between populations revealed significant differentiation between European populations. Both LV and LT were statistically different from CEU at the rs2277460 ($P < 0.05$ and $P < 0.01$ respectively). LT differed from both CEU and LV at the rs2295826 and rs2295827 ($P < 0.05$) and from LV at the rs1048990 ($P < 0.01$). No differentiation was revealed between TW, HCB and JPT. All LV, LT and CEU strongly ($P < 0.001$) differed from Asians at the rs2277460, rs1048990 and rs2348071. However, at the rs2295826 and rs2295827 Asians were different only from LV and CEU; and exhibited similarities with LT. When five loci (1–5) haplotypes were taken into consideration, pairwise differentiation between Europeans and Asians was of strong significance ($P < 0.001$) and LV was different from LT at nominal level of significance ($P < 0.05$). Finally, we estimated population differentiation for *PSMA6* (rs2277460-rs10489990) and *PSMC6* (rs2295826-rs2295827) two locus haplotypes and obtained significantly different patterns of differentiation. The *PSMA6* haplotype differentiated all European populations from Asians ($P < 0.001$ for all pairs) and LT from both LV and CEU ($P < 0.05$) similar to five locus haplotype. When considering the *PSMC6* haplotype, LV and CEU were differentiated from LT ($P < 0.05$) and all Asian populations ($P < 0.001$); LT was undifferentiated from Asians ($P > 0.05$).

Eventual functional significance of nucleotide substitutions

Figs. 2–4 illustrate our findings on eventual functional significance of nucleotide substitutions evaluated as loss/generation of TFBSs and splicing signals in the *PSMA6*, *PSMC6* and *PSMA3* genes respectively.

Table 4
Pairwise estimates of F_{st} between human populations using individual SNP (1, 3, and 5 — above first three diagonals and 2, 4 — below first two diagonals), 5 SNPs (1–5 haplotype — below third diagonal), *PSMA6* (1–2 haplotype — above forth diagonal) and *PSMC6* (3–4 haplotype — below forth diagonal) SNPs frequencies.

	SNP	LV	LT	TW	CEU	HCB	JPT
LV	1	–	0.00228	0.04483***	0.03655*	0.06299*	0.06299*
LT	2	0.01623**	–	0.07169***	0.06317**	0.09030**	0.09030**
TW		0.15592***	0.07675***	–	−0.00592	0.00684	0.00684
CEU		−0.00291	−0.00097	0.11528***	–	0.0000	0.0000
HCB		0.21074***	0.12179***	0.00162	0.16684***	–	0.0000
JPT		0.23151***	0.14021***	0.00794	0.18688***	−0.01141	–
LV	3	–	0.01482*	0.00958*	0.00504	0.01850*	0.01822*
LT	4	0.01482*	–	−0.00107	0.04788*	−0.00713	−0.00682
TW		0.00664*	0.00004	–	0.03787*	−0.00317	−0.00299
CEU		0.00504	0.04788*	0.03220*	–	0.05595*	0.05534*
HCB		0.01850*	−0.0071	−0.00117	0.05595*	–	−0.01221
JPT		0.01822*	−0.00682	−0.00106	0.05534*	−0.01221	–
LV	5	–	−0.00122	0.22986***	−0.00658	0.26355***	0.24855***
LT	1–5	0.00838*	–	0.26109***	−0.00607	0.29571***	0.28040***
TW		0.13338***	0.11657***	–	0.23385***	−0.00385	−0.00509
CEU		0.00107	0.02075**	0.14034***	–	0.26784***	0.25274***
HCB		0.16164***	0.13569***	−0.00145	0.17120***	–	−0.011181
JPT		0.16205***	0.13476***	0.00023	0.17074***	−0.01184	–
D							
LV	1–2	–	0.01058*	0.13730***	0.00843	0.19030***	0.20896***
LT	3–4	0.01482*	–	0.07570***	0.01847*	0.11609***	0.13140***
TW		0.00813*	−0.00052	–	0.10967***	0.00170	0.0792
CEU		0.00504	0.04788*	0.03508	–	0.16266***	0.18235***
HCB		0.01850*	−0.00713	−0.00218	0.05595*	–	−0.01141
JPT		0.01822*	−0.00682	−0.00204	0.05534*	−0.01221	–

* indicates F_{st} values significantly different from zero: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

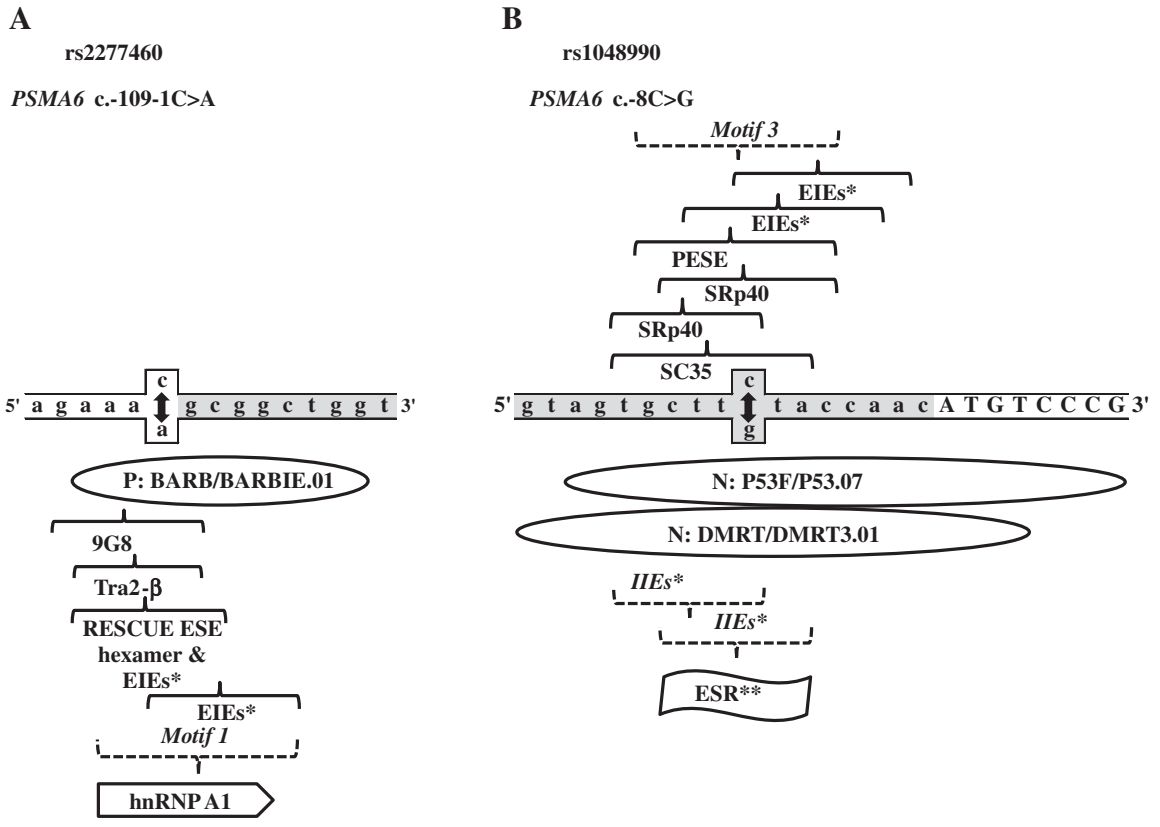

Fig. 2. Consequences of the rs2277460 (Panel A) and rs1048990 (Panel B) nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMA6* gene. Promoter and exon are coloured in white, 5′-UTR is coloured in grey; sequences of coding and noncoding genes' regions are presented by capital and small letters respectively. Positive and negative DNA strands are indicated by capital letters P and N respectively. The transcription factors family and matrix names are separated by symbol of division and given according to MatInspector, Release 7.4 online tool at www.genomatix.de/: BARB/BARBIE.01 — barbiturate-inducible element; P53F/P53.07 — tumour suppressor p53; DMRT/DMRT3.01 — double sex and mab-3 related TF 3. Splicing enhancers are indicated by solid up-directed horizontal braces; splicing silencers are indicated by interrupted down-directed horizontal braces; splicing enhancer and silencers motifs are abbreviated according to Human Splicing Finder Version 2.4 at http://www.umd.be/HSF. Other abbreviations: ESR — exonic splicing regulatory sequence. Asterix (*) indicates situation when several splicing signals of the same type could occupy the sequence and overlap each other.

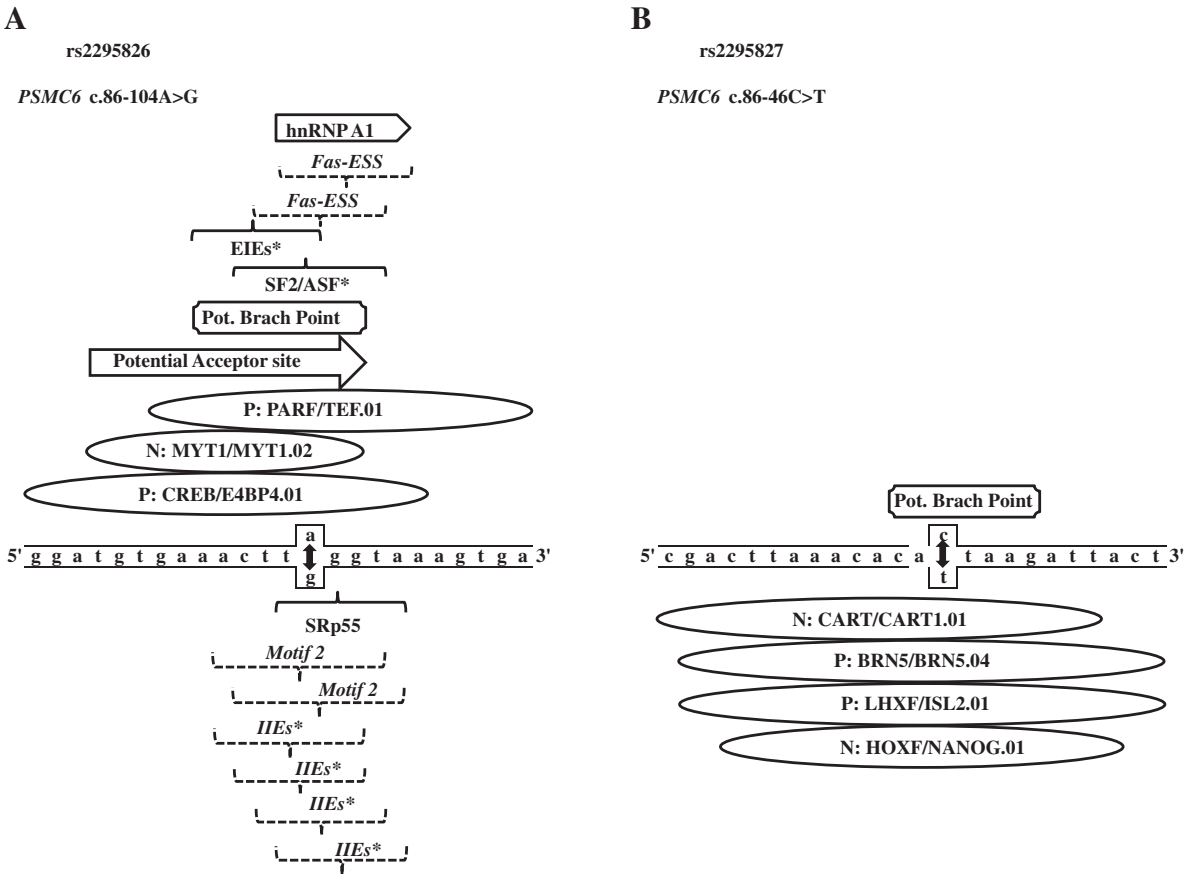

Fig. 3. Consequences of the rs2295826 (Panel A) and rs2295827 (Panel B) nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMC6* gene. Sequence features and other marks are given as described in Fig. 2. CREB/E4BP4.01–E4BP4, bZIP domain, transcription repressor; MYT1/MYT1.02–MyT1 zinc finger TF involved in primary neurogenesis; PARF/TEF.01–thyrotrophic embryonic factor; CART/CART1.01–Cart-1 cartilage homeoprotein 1; BRN5/BRN5.04–POU class 6 homeobox 1 (POU6F1); LHXF/ISL2.01–ISL LIM homeobox 2; HOXF/NANOG.01–Homeobox TF Nanog. "Pot. Branch Point" means potential branch point.

No TFBSs and splicing signals were predicted for genomic region having the rs2277460 major and ancestral allele C. Substitution for rare allele A, potentially assists in creation of binding site to barbiturate inducible element BARBIE.01, and similarity to number of splicing signals including the hnRNP A1 motif.

Genomic region having the rs1048990 major and ancestral allele C, potentially possess similarity with number of splicing signals; no TFBSs were predicted. Nucleotide substitution to minor G allele significantly change patterns of sequence similarity to splicing signals, assists to sequence similarity with exonic splicing regulatory sequence and creates TFBSs of tumour suppressor p53 and DMRT families.

The rs2295826 major and ancestral allele A makes encompassing sequence to be similar to number of splicing signals including additional branch point, potential acceptor site, the hnRNP A1 motif and number of splicing enhancers and silencers as well as BSs to transcription factors of CREB, MYT1 and PARF families. Substitution to G allele appears to eliminate mentioned activities and makes sequence similar mostly to splicing silencers.

The rs2295827 major and ancestral allele C could generate additional branch point, but the sequence affinity to number of TFs of CART, BRNS, LHXF and HOXF families and similarity to number of splicing signals, depend on the presence of minor T allele.

In contrast to the loci described above, the rs2348071 ancestral allele A is the major allele only in Asian populations being the minor allele in Caucasians. When allele G is present, sequence manifests more similarity to number of splicing signals than in the case of allele A; in turn the allele A assists to sequence similarity with hnRNP A1 motif and creates BSs to TFs of CART, MEF2 and HBOX families.

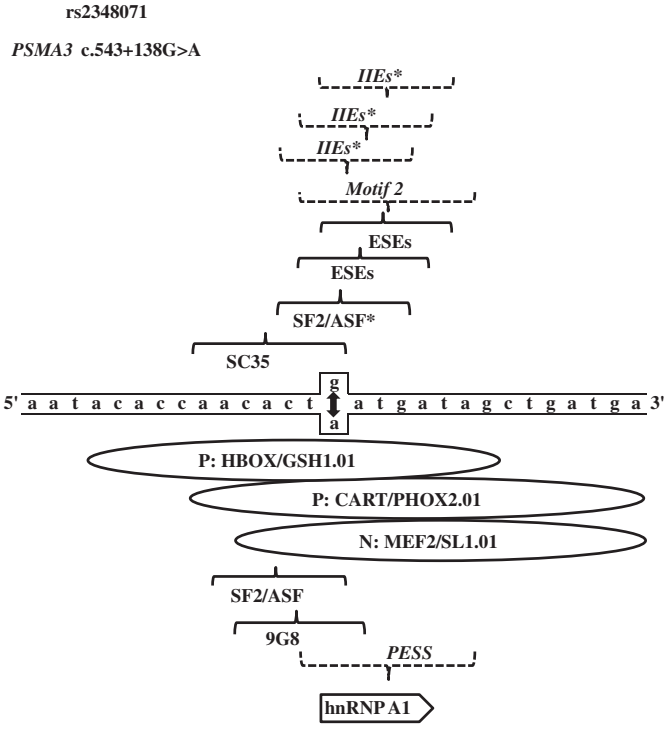

Fig. 4. Consequences of the rs2348071 nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMA3* gene. Sequence features and other marks are given as described in Fig. 2. HBOX/GSH1.01–Homeobox TF Gsh-1; Cart/PHOX2.01–Phox2a (ARIX) and Phox2b of cartilage homeoproteins family; MEF2/SL1.01–member of the RSRF related to serum response factors. Other abbreviations are the same as in Figs. 2 and 3.

Discussion/conclusion

In the current study we investigated genetic polymorphism of the *PSMA6* (rs2277460 and rs1048990), *PSMC6* (rs2295826 and rs2295827) and *PSMA3* (rs2348071) proteasomal genes in 1438 unrelated subjects from LV, LT and TW populations.

The rs2277460 polymorphism locates in the promoter of the *PSMA6* gene in distance of one and 110 nucleotides from the 5′-UTR and translation start respectively. Ancestral allele C, the major allele in human populations over the world, appears to be functionally neutral. Substitution to A generates binding sites to the BARBIE box proteins shown to be involved in signal transduction pathways during development (Arbouzova et al., 2006) and modulation of innate immunity (Dozmorov et al., 2009), hnRNP A1 known as alternative splicing repressor (Clower et al., 2010) and factor facilitating processing of specific microRNAs (Michlewski et al., 2010) as well as number of other splicing signals.

Transversion C → A appears to be rather recent mutation happed and fixed in Europeans in historical time. Alleles' and genotypes' frequencies observed in LV and LT are similar to reported previously for Latvia (Sjakste et al., 2007b, in press; Trapina et al., 2009) and UK (Sjakste et al., 2007b). Rare allele A not detected previously in Asian and African HapMap populations was identified in Taiwanese. In fact, our TW group represented by approximately 50 times more subjects than each HCB and JPT populations and potentially is more informative with respect to genetic diversity and rare allele, genotype and haplotype identification.

The rs1048990 SNP located in 8 nucleotides distance upstream the *PSMA6* initiation codon and four nucleotides upstream the Kozak consensus (Kozak, 1997), potentially could interfere splicing and initiation of translation. In fact, Ozaki et al. (2006) reported that the rs1048990 G allele was associated with higher expression of the *PSMA6* gene *in vitro* and *in vivo*. Recently, Wang et al. (2013) compared levels of mRNA expression by the rs1048990 genotypes and significant trend by G allele was also found. We showed here that both C and G alleles could affect sequence functional potential. Ancestral allele C generates presumably splicing enhancer motifs; minor allele G creates an exonic splicing regulatory motif, several splicing silencers and binding sites for tumour suppressor p53 and double sex and mab-3 related transcription factor of DMRT family. Gene encoding the tumour suppressor p53 was shown to express ethnic heterogeneity and may be involved in ecological (climatic) adaptation (Själander et al., 1996). TFs of the DMRT family were shown to play significant roles during animal evolution contributing to the origin of novel sex-specific traits (Kopp, 2012) and potentially should be involved in human adaptation.

Ancestral allele C is a major allele in all human populations over the world being significantly less frequent in Asians than in Europeans (about 70% vs 90% in HCB and CEU populations respectively). Similarly, in our study, we have observed the minor allele G approximately four and twice more frequent in TW compared to LV and LT respectively. Transversion C → G is one of the oldest mutations studied here and appears to happen approximately 8000–11,000 years ago in Asians. In Europeans this mutation appears to arise much later evolutionally or, more probably, had been introduced with human migrations already in historical time.

Multiple case/control studies conducted during the last decade to search for association between the rs1048990 SNP and human diseases provide significant information on locus variability. Firstly, different studies applied for the same ethnic groups (Latvians Sjakste et al., 2007b, in press; Trapina et al., 2009), British (Bennett et al., 2008; Freilinger et al., 2009; Sjakste et al., 2007b), Indians (Banerjee et al., 2008, 2009) and Japanese (Hinohara et al., 2009; Ikeda et al., 2012; Ozaki et al., 2006; Takashima et al., 2007) showed similar allele and genotype frequencies suggesting that control cohorts of case/control studies could successfully represent corresponding population. Secondly, locus variability appears to express geographical- and/or ethnos-specific dynamic. Minor G allele appears to be least of all presented in YRI (MAF of 0.017). In Europe, MAF appears to increase from North East (mean for LV from current and previous (Sjakste et al., 2007b, in press; Trapina et al., 2009) studies equal to 0.107) to South West (0.133 in Ukraine (Honcharov et al., 2009), mean of 0.159 for UK (Bennett et al., 2008; Freilinger et al., 2009; Ozaki et al., 2006) and 0.169 in Germany (Freilinger et al., 2009)) remaining significantly rarer than in India (mean of 0.206 Banerjee et al., 2008, 2009) and Saudi Arabia (0.231 Alsmadi et al., 2009). Allele G showed the highest frequency in Eastern Asians where MAF varies from 0.317 in Japanese (mean from Hinohara et al., 2009; Ikeda et al., 2012; Takashima et al., 2007) and 0.320 in Chinese (Liu et al., 2009) till 0.350 in Koreans (Hinohara et al., 2009).

In LV and LT groups of the current study and in vast majority of populations of published studies, the rs1048990 SNP was found to occur in frequencies consistent with HWE. However, genotype distribution was found to deviate significantly from HWE in our TW population, South Italian (Barbieri et al., 2008), Saudi (Alsmadi et al., 2009) and one of Japanese control cohort (Ikeda et al., 2012). Three statistics applied in our study did not show significant deviation from neutral model of evolution.

The rs1048990 locus susceptibility reported for several pathologies (Table 1) appears to have ethnos-specific character (Wang et al., 2013). We suggest that ethnos- and/or geographically-specific allele and genotype distribution at the rs1048990 is an evolutionary natural phenomenon involved originally in the mechanisms of ethnic adaptation to the definite environment and may influence general morbidity of human populations.

The rs2295826 and rs2295827 both located in the first intron of the *PSMC6* gene in 61 bp from each other, showed an r^2 between 0.923 in Tuscans (Italy) and 1.0 (CEU) in different Caucasian ethnicities and a D′ of 1.0 in all ethnicities analysed until the current study, suggesting three AC, GT and GC the rs2295826/rs2295827 haplotypes. In both LV and LT populations, the rs2295826 and rs2295827 showed the same alleles' and genotypes' frequencies suggesting strong linkage between the loci and only AC and GT haplotypes' occurrence. To our surprise, we did not observed in Taiwanese the rs2295827 rare allele TT homozygotes. This fact suggests a disruption of linkage between the rs2295826 and rs2295827 loci ($D = 0.978$; $r^2 = 0.901$) and occurrence of forth rare AT haplotype in Taiwanese.

The rs2295826 A and rs2295827 C ancestral alleles were the major in all populations analysed. Alleles and genotypes distributions were similar between LV and CEU and TW and Asian HapMap populations respectively. To our surprise, minor alleles were significantly more frequent in LT than in LV and CEU.

The major allele of the rs2295826 generates an additional splice site acceptor and branch point, hnRNP A1 and several splicing enhancer and silencer motifs as well as sequence affinity to TFs of CREB, MYT1 and PARF families known to be involved, in regulation of multiple physiological processes including control of circadian clock (Male et al., 2012; Wang et al., 2010). Genetic variation occurred within coding and non-coding regions of several genes regulating circadian rhythm was shown to be ethnos specific (Cruciani et al., 2008; Hawkins et al., 2008) and might represent an evolutional history of adaptation in populations of different geographic origin.

The rs2295826 minor allele G generates splicing silencers mostly. Additional branch point is predicted for sequence encompassing the rs2295827 major allele C. Sequences having the rs2295827 minor allele G can potentially bind the CART proteins responsible for bone and cartilage development (Furukawa et al., 2002), BRN5 and LHXF factors known to mediate transcriptional control of neuronal differentiation (Gill, 2003; Phillips and Luisi, 2000; She and Mao, 2011; Uzumcu et al., 2009) and HOXF family NANOG.01 factor shown to be generally involved in signal transduction pathways during development (Ho et al., 2012).

The rs2348071 polymorphism located in the intron 7 of the *PSMA3* gene strongly discriminates Asians having a major ancestral allele A (about 70%) and other ethnics having a major allele G (about 70%). Transition A → G appears to be one of the oldest among analysed mutations which happened in Caucasians about 15,000 years ago and was supported by positive selection in Caucasians over the world. Mutation age appears to be less in Asians and might result from both the *de novo* mutation event and gene flow from other ethnics.

Similar to the rs2348071, Wang et al. (2008) identified several loci in the *PSMB1*, *PSMB2* and *PSMB5* proteasomal genes being observed as minor in one ethnic group and middle/common in others and suggested that clinical response to proteasomal inhibitors potentially might be allele specific (hypothesis is discussed in Wang et al. (2008)).

The above mentioned and our findings indicate that loci greatly diverse between the populations in the allele and genotype presentation could potentially be involved in processes of evolutional and/or geographical adaptation of human populations to the environment. These should be of special interest and perspective for medical applications.

The rs2348071 ancestral allele A generates binding sites for already mentioned CART proteins and MEF2 and HBOX factors known to mediate transcriptional control of neuronal differentiation (Gill, 2003; Phillips and Luisi, 2000; She and Mao, 2011;Uzumcu et al., 2009) and generates splicing signals including the hnRNP A1 and several enhancer and silencer motifs. Sequence having G allele appears to have a big potential in respect of splicing regulation as many different splicing signals might be involved in regulation.

So, allele specific sequence functional motifs potentially could significantly affect particular gene expression, UPS functionality in general and network of different genes and proteins including those involved in ethnic specific adaptation to environment.

As it was expected, Taiwanese were significantly different from both the LV and LT population in the rs2277460, rs1048990 and the rs2348071 genetic diversity. However, at the rs2295826 and rs2295827, TW was different only from LV and exhibited similarity with LT. LT and LV populations showed different genetic diversity at the rs1048990, the *PSMC6* SNPs, the *PSMA6* rs2277460-rs10489990 haplotype and the 5-locus haplotype.

Finding of differentiation between the LT and LV populations appears not to be too unexpected. Living geographically close to each other and in big extent share a common paternal Y chromosome and maternal mitochondrial gene pools (Kasperaviciute et al., 2004; Laitinen et al., 2002; Lessig et al., 2001; Pliss et al., 2006), LV and LT population were shown to be different in particular markers' frequency and haplogroups' diversity. Examples include the Y chromosome M9 rare allele C and haplogroup HG2 being twice less frequent in LV and haplogroup HG1 observed twice more frequent in LV than in LT (Laitinen et al., 2002). Analysis of 270,000 SNPs genotyped in samples of DNA collected all over Europe revealed that the genetic structure of the European population correlates closely with geography and markers are grouped in a less compact way in Latvian population compared to Lithuanian population where more similarities with Central Europe were revealed (Nelis et al., 2009). Allele and genotype presentation of the rs1048990 was found also to be more similar in LT with that observed in Central Europe. We did not perform ethno genetic study; ethnic origin of subjects was not recorded during sample collection. However evidently LV population in our case represents very mixed inhabitants of Riga, forming some “average” genotype for North-East Europe. On the contrary inhabitants of Kaunas are mostly Lithuanians, population is ethnically homogenous. Admixture of non-Baltic ethnic groups in Riga is reflected as similarity to the European population in general, ethnic peculiarities were revealed for Lithuania.

Similarities between LT population and Asians seem even more striking at the first glance. It is considered that Ural mountains and Caspian sea form a border between Asian and European populations (Kutuev et al., 2006), however numerous migrations across this border could result in genetic material transfer, as it was shown for some Y chromosome haplogroups (Rootsi et al., 2007). Moreover similar distribution of the D7S23 locus allelic frequencies was found among LT population, Bashkirs inhabiting Volga region, Komi living on the North-East border of Europe, and Buryat population from East Siberia (Khusnutdinova et al., 1994). Invasions of Asian peoples to Baltics happened several times in history, some distinct groups of Turkic-speaking peoples still live in Lithuania. Thus, geographically close Baltic populations might have different microevolution of human genome acting on different traits and metabolic pathways including ubiquitin proteasome system.

In conclusion, we suggest that ethnic- and/or geographically-specific patterns of structural variations in proteasomal genes may reflect processes of local historical and/or geographical ethnos adaptation and reserve influence for human health and population morbidity in modern environment.

Acknowledgements

Costs of this work were partly covered by the following projects: Latvia–Lithuania–Taiwan Cooperation no V320100512; Latvian National Research Programme 2010.10.-4/VPP4; and ESF project No 2013/0043/1DP/1.1.1.2.0/13/APIA/VIAA/002. The Lithuanian team acknowledges the Ministry of Education and Science and Research Council of Republic of Lithuania for financial agreements TAP 10043, TAP-06/2011 and TAP LLT 08/2012.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.mgene.2014.03.002.

References

Alsmadi, O., Muiya, P., Khalak, H., et al., 2009. Haplotypes encompassing the KIAA0391 and PSMA6 gene cluster confer a genetic link for myocardial infarction and coronary artery disease. Ann. Hum. Genet. 73, 475–483.

Arbouzova, N.I., Bach, E.A., Zeidler, M.P., 2006. Ken & Barbie selectively regulates the expression of a subset of Jak/STAT pathway target genes. Curr. Biol. 16, 80–88.
Bachmann, H.S., Novotny, J., Sixt, S., et al., 2010. The G-allele of the PSMA6-8G>C polymorphism is associated with poor outcome in multiple myeloma independently of circulating proteasome serum levels. Eur. J. Haematol. 85, 108–113.
Banerjee, I., Gupta, V., Ahmed, T., et al., 2008. Inflammatory system gene polymorphism and the risk of stroke: a case–control study in an Indian population. Brain Res. Bull. 75, 158–165.
Banerjee, I., Pandey, U., Hasan, O.M., et al., 2009. Association between inflammatory gene polymorphisms and coronary artery disease in an Indian population. J. Thromb. Thrombolysis 27, 88–94.
Barbieri, M., Marfella, R., Rizzo, M.R., et al., 2008. The −8 UTR C/G polymorphism of PSMA6 gene is associated with susceptibility to myocardial infarction in type 2 diabetic patients. Atherosclerosis 201, 117–123.
Bedford, L., Lowe, J., Dick, L.R., et al., 2011. Ubiquiti-like protein conjugation and the ubiquitin-proteasome system as drug targets. Nat. Rev. 10, 29–46.
Bennett, D.A., Xu, P., Clarke, R., et al., 2008. International Study of Infarct Survival Collaborators. The exon 1-8C/G SNP in the PSMA6 gene contributes only a small amount to the burden of myocardial infarction in 6946 cases and 2720 controls from a United Kingdom population. Eur. J. Hum. Genet. 16, 480–486.
Cartharius, K., Frech, K., Grote, K., 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 21, 2933–2942.
Clower, C.V., Chatterjee, D., Wang, Z., et al., 2010. The alternative splicing repressors hnRNP A1/A2 and PTB influence pyruvate kinase isoform expression and cell metabolism. Proc. Natl. Acad. Sci. U. S. A. 107, 1894–1899.
Cruciani, F., Trombetta, B., Labuda, D., et al., 2008. Genetic diversity patterns at the human clock gene period 2 are suggestive of population-specific positive selection. Eur. J. Hum. Genet. 16, 1526–1534.
Desmet, F.O., Hamroun, D., Lalande, M., et al., 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 37, e67.
Dozmorov, M., Wu, W., Chakrabarty, K., et al., 2009. Gene expression profiling of human alveolar macrophages infected by *B. anthracis* spores demonstrates TNF-alpha and NF-kappab are key components of the innate immune response to the pathogen. BMC Infect. Dis. 9, 152.
Freilinger, T., Bevan, S., Ripke, S., et al., 2009. Genetic variation in the lymphotoxin-alpha pathway and the risk of ischemic stroke in European populations. Stroke 40, 970–972.
Fu, Y.-X., Li, W.-H., 1993. Statistical tests of neutrality of mutations. Genetics 133, 693–709.
Furukawa, K., Iioka, T., Morishita, M., et al., 2002. Functional domains of paired-like homeoprotein Cart1 and the relationship between dimerization and transcription activity. Genes Cells 7, 1135–1147.
Gill, G.N., 2003. Decoding the LIM development code. Trans. Am. Clin. Climatol. Assoc. 114, 179–189.
Hawkins, G.A., Meyers, D.A., Bleecker, E.R., et al., 2008. Identification of coding polymorphisms in human circadian rhythm genes PER1, PER2, PER3, CLOCK, ARNTL, CRY1, CRY2 and TIMELESS in a multi-ethnic screening panel. DNA Seq. 19, 44–49.
Heckman, M.G., Soto-Ortolaza, A.I., Diehl, N.N., et al., 2013. Genetic variants associated with myocardial infarction in the PSMA6 gene and Chr9p21 are also associated with ischaemic stroke. Eur. J. Neurol. 20, 300–308.
Hinohara, K., Nakajima, T., Sasaoka, T., et al., 2009. Replication studies for the association of PSMA6 polymorphism with coronary artery disease in East Asian populations. J. Hum. Genet. 54, 248–251.
Ho, B., Olson, G., Figel, S., et al., 2012. Nanog increases focal adhesion kinase (FAK) promoter activity and expression and directly binds to FAK protein to be phosphorylated. J. Biol. Chem. 287, 18656–18673.
Honcharov, S.V., Dosenko, V.I., Khaitovych, M.V., et al., 2009. Allele polymorphism of genes coding proteasome subunits is associated with an enhanced risk for arterial hypertension in adolescents. Fiziol. Zh. 55, 3–10.
Ikeda, S., Tanaka, N., Arai, T., et al., 2012. Polymorphisms of LTA, LGALS2, and PSMA6 genes and coronary atherosclerosis: a pathological study of 1503 consecutive autopsy cases. Atherosclerosis 221, 458–460.
Kalis, M., Sjakste, T., Sjakste, N., et al., 2002. Association study between (TG) repeat polymorphism in PSMA6 gene and type II dibetes mellitus in Botnia. Biologija 2, 12–14.
Kasperaviciute, D., Kucinskas, V., Stoneking, M., 2004. Y chromosome and mitochondrial DNA variation in Lithuanians. Ann. Hum. Genet. 68, 438–452.
Khusnutdinova, E.K., Khidiatova, I.M., Ivashcenko, T.E., et al., 1994. Polymorphism of MET and D7S23 loci linked to the cystic fibrosis gene in Bashkir and Komi populations. Hum. Hered. 44, 191–194.
Kopp, A., 2012. Dmrt genes in the development and evolution of sexual dimorphism. Trends Genet. 28, 175–184.
Kozak, M., 1997. Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. EMBO J. 16, 2482–2492.
Kupca, S., Sjakste, T., Paramonova, N., et al., 2013. Association of obesity with proteasomal gene polymorphisms in children. J. Obes. 2013, 638154.
Kutuev, I., Khusainova, R., Karunas, A., et al., 2006. From East to West: patterns of genetic diversity of populations living in four Eurasian regions. Hum. Hered. 61, 1–9.
Laitinen, V., Lahermo, P., Sistonen, P., et al., 2002. Y-chromosomal diversity suggests that Baltic males share common Finno-Ugric-speaking forefathers. Hum. Hered. 53, 68–78.
Lessig, R., Edelmann, J., Krawczak, M., 2001. Population genetics of Y-chromosomal microsatellites in Baltic males. Forensic Sci. Int. 118, 153–157.
Liu, X., Wang, X., Shen, Y., et al., 2009. The functional variant rs1048990 in PSMA6 is associated with susceptibility to myocardial infarction in a Chinese population. Atherosclerosis 206, 199–203.
Male, V., Nisoli, I., Gascoyne, D.M., 2012. Brady HJ. E4BP4: an unexpected player in the immune response. Trends Immunol. 33, 98–102.
Michlewski, G., Guil, S., Cáceres, J.F., 2010. Stimulation of pri-miR-18a processing by hnRNP A1. Adv. Exp. Med. Biol. 700, 28–35.
Nelis, M., Esko, T., Mägi, R., et al., 2009. Genetic structure of Europeans: a view from the North-East. PLoS One 4, e5472.
Ozaki, K., Sato, H., Iida, A., et al., 2006. A functional SNP in PSMA6 confers risk of myocardial infarction in the Japanese population. Nat. Genet. 38, 921–925.
Phillips, K., Luisi, B., 2000. The virtuoso of versatility: POU proteins that flex to fit. J. Mol. Biol. 302, 1023–1039.

Pliss, L., Tambets, K., Loogväli, E.L., et al., 2006. Mitochondrial DNA portrait of Latvians: towards the understanding of the genetic structure of Baltic-speaking populations. Ann. Hum. Genet. 70, 439–458.
Rodriguez, S., Gaunt, T.R., Day, I.N., 2009. Hardy–Weinberg equilibrium testing of biological ascertainment for Mendelian randomization studies. Am. J. Epidemiol. 169, 505–514.
Rootsi, S., Zhivotovsky, L.A., Baldovic, M., et al., 2007. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. Eur. J. Hum. Genet. 15, 204–211.
Rozas, J., 2009. DNA sequence polymorphism analysis using DnaSP. Bioinformatics for DNA sequence analysis; methods. Mol. Biol. Ser. 537, 337–350.
She, H., Mao, Z., 2011. Regulation of myocyte enhancer factor-2 transcription factors by neurotoxins. Neurotoxicology 32, 563–566.
Sjakste, T., Eglite, J., Sochnev, A., et al., 2004. Microsatellite genotyping of chromosome 14q13.2-14q13 in the vicinity of proteasomal gene PSMA6 and association with Graves' disease in the Latvian population. Immunogenetics 56, 238–243.
Sjakste, T., Kalis, M., Poudziunas, I., et al., 2007a. Association of microsatellite polymorphisms of the human 14q13.2 region with type 2 diabetes mellitus in Latvian and Finnish populations. Ann. Hum. Genet. 71, 772–776.
Sjakste, T., Poudziunas, I., Ninio, E., et al., 2007b. SNPs of PSMA6 gene-investigation of possible association with myocardial infarction and type 2 diabetes mellitus. Genetika 43, 553–559.
Sjakste, T., Trapina, I., Rumba-Rozenfelde, I., et al., 2010. Identification of a novel candidate locus for juvenile idiopathic arthritis at 14q13.2 in the Latvian population by association analysis with microsatellite markers. DNA Cell Biol. 29, 543–551.
Sjakste, T., Paramonova, N., Rumba-Rozenfelde, I., et al., 2014. Juvenile idiopathic arthritis subtype and sex specific associations with genetic variants in the *PSMA6*/*PSMA3*/*PSMC6* gene cluster. Pediatr. Neopathol. (in press).
Själander, A., Birgander, R., Saha, N., et al., 1996. P53 polymorphisms and haplotypes show distinct differences between major ethnic groups. Hum. Hered. 46, 41–48.
Sorokin, A.V., Kim, E.R., Ovchinnikov, L.P., 2009. Proteasome system of protein degradation and processing. Biochemistry (Moskow) 74, 1411–1442.
Tajima, F., 1989. Statistical method for testing the neutral mutationhypothesis by DNA polymorphism. Genetics 123, 585–595.
Takashima, N., Shioji, K., Kokubo, Y., et al., 2007. Validation of the association between the gene encoding proteasome subunit α type 6 and myocardial infarction in a Japanese population. Circ. J. 71, 495–498.
Trapina, I., Rumba-Rozenfelde, I., Sjakste, N., et al., 2009. Association study of genetic variants in the 14q11–14q13 proteasomal genes cluster with juvenile idiopathic arthritis (JIA) in Latvian population. Proc. Latv. Acad. Sci. B. 63, 20–30.
Uzumcu, A., Karaman, B., Toksoy, G., et al., 2009. Molecular genetic screening of MBS1 locus on chromosome 13 for microdeletions and exclusion of FGF9, GSH1 and CDX2 as causative genes in patients with Moebius syndrome. Eur. J. Med. Genet. 52, 315–320.
Wang, L., Kumar, S., Fridley, B.L., 2008. Proteasome beta subunit pharmacogenomics: gene resequencing and functional genomics. Clin. Cancer Res. 14, 3503–3513.
Wang, Q., Maillard, M., Schibler, U., et al., 2010. Cardiac hypertrophy, low blood pressure, and low aldosterone levels in mice devoid of the three circadian PAR bZip transcription factors DBP, HLF, and TEF. Am. J. Physiol. Regul. Integr. Comp. Physiol. 299, R1013–R1019.
Wang, H., Jiang, M., Zhu, H., et al., 2013. Quantitative assessment of the influence of PSMA6 variant (rs1048990) on coronary artery disease risk. Mol. Biol. Rep. 40, 1035–1041.
Willis, M.S., Townley-Tilson, W.H.D., Kang, E.Y., et al., 2010. Send to destroy: the ubiquitin proteasome system regulates cell signalling and protein quality control in cardiovascular development and disease. Circ. Res. 106, 463–478.
Zemeckienė, Z., Vitkauskienė, A., Sjakste, T., et al., 2013. Proteasomes and proteasomal gene polymorphism in association with inflammation and various diseases. Medicina (Kaunas) 49, 207–213.