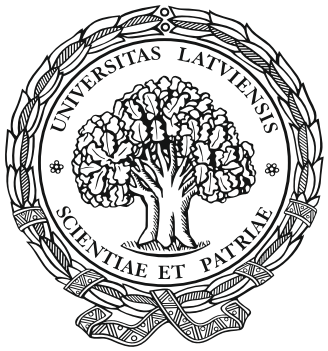
LATVIAN UNIVERSITY
FACULTY OF BIOLOGY



Natalia Paramonova

POLYMORPHISMS IN PROTEASOME GENE CLUSTER OF HUMAN CHROMOSOME 14 AND THEIR ASSOCIATION WITH AUTOIMMUNE AND METABOLIC DISEASES IN LATVIAN POPULATION

DOCTORAL THESIS
Submitted for the Doctoral Degree in Biology
Subfield of Molecular Biology

RIGA, 2015

Supervisor:

***Dr. habil. biol*, prof. Nikolajs Sjakste**

Reviewers:

Prof. J. Rzeszowska-Wolny, Silesian Technical University, Gliwice, Poland
Dr med., Liene Ņikitina-Zaķe, Latvian Biomedical Research and Study Centre
Dr med., prof. Valdis Pīrāgs, Latvian University, Faculty of Medicine Department of Inner Diseases

Chairman of the Doctoral Commitee

Dr med., prof. Valdis Pīrāgs

Secretary of the Doctoral Commitee

Dr pharm., Kristīne Saleniece

This work has been supported by the Eiropean Social Foundation Project 2013/0043/1DP/1.1.1.2.0/13/APIA/VIAA/002

IEGULDĪJUMS TAVĀ NĀKOTNĒ

SUMMARY

Autoimmune and metabolic diseases cause numerous health and social problems throughout the world. The possible association of polymorphisms in Ubiquitin proteasomes system (UPS) subunits with human pathologies is intensively sought. We have focused our attention on juvenile idiopathic arthritic (JIA), bronchial asthma (BA) and obesity (OB). The candidate genes from proteasomes genes clusters located in the 14q11.2 - 14q23 chromosomal region, earlier investigated as diseases susceptible, was chosen for the current association study.

Goal of the study was to investigate genetic diversity of proteasomes genes cluster located in 14q11.2 - 14q23 chromosomal region in Latvian (LV), Lithuanian (LT), and Taiwanese (TW) populations and to reveal possible associations of investigated loci with BA in LV and TW; with JIA, BA and OB in Latvian population.

The significant differences were revealed in distribution of *PSMA6* (rs2277460) and *PSMC6* (rs2295826 and rs2295827) proteasomal genes variations in Latvian, Lithuanian and Taiwanese populations. Both LV, LT strongly differed from TW at the rs2277460, rs1048990 and rs2348071. However, at the rs2295826 and rs2295827 Taiwanese were different only from LV and exhibited similarities with LT (Publication I).

Our results suggest an association of the 14q13-23 proteasomal genes polymorphisms with childhood asthma in Latvians and Taiwanese and highlight risk and/or protective factors being the same or different between the populations (Publication III).

Variations at the *PSMA6* (rs2277460 and rs1048990), *PSMC6* (rs2295826 and rs2295827) and *PSMA3* (rs2348071) contribute to JIA susceptibility; the *PSMA6/PSMC6/PSMA3* genetic variants and multiloci genetic modules could be suggested as JIA subtype- and sex-specific risk factors (Publication II).

We demonstrate for the first time evidence of a sex-specific association of *PSMA6/PSMC6/PSMA3* genetic variants with subtypes of JIA and plasma proteasome concentrations (Publication II).

Variations at the *PSMB5* (rs11543947) and *PSMC6* (rs2295826 and rs2295827), manifested association with obesity in total diseases group and could be suggested as OB familial and sex-specific risk factors (Publication IV).

In silico analysis reveals that that the nucleotide substitutions we have studied modify transcription factor binding sites and miRNAs, this can significantly modulate the transcription of related genes and gene network in response to the inflammation and other environmental stimuli and influence the diseases susceptibility (Publication I, II).

Table of contents

ABBREVIATIONS

14ch	Chromosome 14
14q	Long arm of Chromosome 14
3'UTR	Gene 3 'untranslated part
5'UTR	Gene 5 'untranslated part
95% CI	Confidence interval 95%
ASA	Allele specific amplification
BA	Bronchial asthma
BMC	Latvian Biomedical Research and Study Centre
BMI	Body weight index (body mass index)
Bp	Base pairs
CAD	Coronary Artery Disease
CAPS	Cleaved amplified polymorphic site.SNP single nucleotide polymorphism
cDNA	Complementary deoxyribonucleic acid
Del	Deletion
DM	Diabetes mellitus
DN	Diabetic kidney disease, or diabetic nephropathy
DNA	Deoxyribonucleic acid
GWAS	Genome-Wide Association Studies database
HLA	Human leukocyte antigen (Human leukocyte antigen)
ILAR	International League of Association for Rheumatology
Ins	Insertion
IS	Ischemic Stroke
JIA	Juvenile Idiopathic arthritis
JIoA	Juvenile idiopathic oligoarthritis
JIpA	Juvenile idiopathic polyarthritis
LV	Latvian population
LT	Lithuanian population
MAF	Rare allele frequency (minor allele frequency)
MI	Myocardial Infarction
MicroRNA	Micro ribonucleic acid
mRNA	Messenger ribonucleic acid
MS	Multiple sclerosis

MSs	Microsatellites
NCBI	National Center of Biotechnology Information (National Center for Biotechnology Information, US; www.ncbi.nlm.nih.gov/)
OB	Obesity
OR	Odds ratio
P	*P*-value, a function of the observed sample results that is used for testing a statistical hypothesis
PCR	Polymerase Chain Reaction
RA	Rheumatoid arthritis
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus
TF	Transcription factor
TFBS	Transcription factor binding site
TW	Taiwanese population
Ub	Ubiquitin
UPS	Ubiquitin proteasomes system

LIST OF PUBLICATIONS AND PRESENTATIONS

The Doctoral Thesis is a summary of publications (articles)

I. Sjakste T, **Paramonova N**, Shi-Shin Wu L, Zemeckiene Z, Sitkauskiene B, Sakalauskas R. Wang J.-Y, Sjakste N. *PSMA6* (rs2277460, rs1048990), *PSMC6* (rs2295826, rs2295827) and *PSMA3* (rs2348071) genetic diversity in Latvians, Lithuanians and Taiwanese. *Meta Gene.* 2014; 2: 283–98.

II. Sjakste T, **Paramonova N**, Rumba-Rozenfelde I, Trapina I, Sugoka O, Sjakste N. Juvenile Idiopathic Arthritis Subtype- and Sex-specific Associations with Genetic Variants in the *PSMA6/PSMC6/PSMA3* Gene Cluster. *Pediatr Neonatol.* 2014; 55(5): 393–403.

III. **Paramonova N,** Shi-Shin Wu L, Rumba-Rozenfelde I, Wang J.-Y, Sjakste N, Sjakste T. Genetic variants in the *PSMA6*, *PSMC6* and *PSMA3* genes associated with childhood asthma in Latvian and Taiwanese populations. *Biopolym. Cell.* 2014; 30(5): 377–87.

IV. **Paramonova N,** Kupca S, Rumba-Rozenfelde I, Sjakste N, Sjakste T. Association between the *PSMB5* and *PSMC6* genetic variations and children obesity in the Latvian population. *Biopolym. Cell.* 2014. 30(6): 477–80.

V. Sjakste T, **Paramonova N,** Sjakste N. Functional significance of microsatellite markers. *Medicina (Kaunas).* 2013; 49(12): 505–9.

Publications in international journals

1. Zemeckiene Z, Sitkauskiene B, Gasiuniene E, **Paramonova N**, Tamasauskiene Zilinskaite L, Vitkauskiene A, Sjakste T, Sakalauskas R. Evaluation of proteasomal gene polymorphisms in Lithuanian patients with asthma. *J Asthma.* 2014; **6**:1–22.
2. Kalnina J, **Paramonova N**, Sjakste N, Sjakste T. Study of association between polymorphisms in the *PSMB5* (rs11543947) and *PSMA3* (rs2348071) genes and multiple sclerosis in Latvians. *Biopolym. Cell.* 2014. 30(4): 305–09.
3. Kupca S, Sjakste T, **Paramonova N**, Sugoka O, Rinkuza I, Trapina I, Daugule I, Sipols A, Rumba-Rozenfelde I. Association of Obesity with Proteasomal Gene Polymorphisms in Children, *J Obes.* 2013. 638154.
4. Wu LS, Sjakste T, Sakalauskas R, Sitkauskiene B, **Paramonova N,** Gasiuniene E, Jan RL, Wang JY. The burden of allergic asthma in children: a landscape comparison based on data from Lithuanian, Latvian, and Taiwanese populations. Pediatr Neonatol. 2012, 53(5): 276–82.

5. Sokolovska J, Isajevs S, Sugoka O, Sharipova J, **Paramonova N**, Isajeva D, Rostoka E, Sjakste T, Kalvinsh I, Sjakste N. Comparison of the Effects of Glibenclamide on Metabolic Parameters, GLUT1 Expression, and Liver Injury in Rats With Severe and Mild Streptozotocin-Induced Diabetes Mellitus. *Medicina (Kaunas),* 2012; 48(10): 532–43.
6. Sjakste T, **Paramonova N**, Grislis Z, Trapina I, Kairisa D.Analysis of the single-nucleotide polymorphism in the 5'UTR and part of intron I of the sheep MSTN gene. *DNA Cell Biol*. 2011, 433–44.

The data are also presented in the congresses and conferences and published is abstracts

1. **Paramonova N.** Sokolovska J, Sjakste T, Sjakste N. variations in the 14q genes encoding proteasome subunits and genetic susceptibility to type 1 diabetes mellitus in Latvians. LU 73. Zinātniskās konferences Medicīnas sekcijas Tēžu apkopojums. Rīga, Latvija, 2015. 20. februāris, lpp. 29.
2. **Paramonova N**, Sjakste T, Rumba-Rozenfelde I, Sjakste N. Association of the *PSMB5* (rs11543947), *PSMC6* (rs2295826, rs2295827) genes polymorphismswith obesity in the Latvian population VII International meeting "From Molecular to Cellular Events in Human Pathologies". Rīga, Latvija, Stenda referāts. 2014, 20. septembris, lpp. 31.
3. Kalnina J, **Paramonova N**, Sjakste N., Sjakste T. Evaluation of the *PSMB5* (rs11543947) and *PSMA3* (rs2348071) gene polymorphisms on the association with multiple sclerosis in Latvians. VII International meeting "From Molecular to Cellular Events in Human Pathologies". Rīga, Latvija, Stenda referāts. 2014, 20. septembris, lpp. 9.
4. **Paramonova N,** Shi-Shin Wu L, Zemeckiene Z, Sakalauskas R, Wang J, Sjakste T. Genetic diversity of proteasomal genes polymorphic loci in human populations. Daugavpils Universitātes 56. Starptautiska zinātniska konference, Šūnas bioloģija, ģenētika un biotehnoloģijas. Daugavpils, Tēžu apkopojums. 2014, 9. - 11. aprīlis, lpp. 23.
5. **Paramonova N**, Sjakste T, Shi-Shin Wu L, Zemeckiene Z, Sugoka O, Sitkauskiene B, Sakalauskas R, Wang J.-Y, Sjakste N. Proteosomu *PSMA6* (rs2277460, rs1048990), *PSMC6* (rs2295826, rs2295827) un *PSMA3* (rs2348071) gēnu ģenētiska daudzveidība Latvijas, Lietuvas un Taivānas populācijās. Latvijas Universitātes 72. konference, Medicīna, Rīga, Latvija, Tēžu apkopojums. 2014, lpp. 30.
6. Sjakste T, **Paramonova N,** Sugoka O, Trapiņa I, Rumba –Rozenfelde I, Sjakste N. Dzimuma un genotipa mijiedarbība proteasomu gēnu polimorfos lokusos. Latvijas

Universitātes 71. konference, Medicīna. Rīga, Latvija. Tēžu apkopojums. 2013, 15. februāris, lpp. 83.

7. Sjakste T, **Paramonova N**, Sugoka O, Trapina I, Rumba-Rozenfelde I, Sjakste N. Proteasomal genes genotype-sex interactions in human populations and in association with complex diseases. Daugavpils Universitātes 55. Starptautiska zinātniska konference, Ģenētika un biotehnoloģija. Daugavpils, Tēžu apkopojums. 2013, 10.-12. aprīlis, lpp. 17.-18.
8. **Paramonova N**, Lunins R, Rumba-Rozenfelde I, Sugoka O, Sjakste N, Sjakste T. A systematic 14q genotyping for association with human pathologies Latvian University 70 Conference, Medicine section, Riga, Latvia, Abstract book. 2012, February 2, pp 83.
9. **Paramonova N**, Trapina N, Sjakste T, Rumba-Rozenfelde I, Lunins R, Sugoka O, Sjakste N. „Jauns 14q13.2 asociācijas lokuss ar autoimūnām slimībām". Latvijas Universitātes ikgadējās konferences ietvaros izstādē "Struktūrfondi Latvijas Universitātē: 2007.-2013", Riga, Latvia. Mutiska postera prezentācija. 2012, 28. februāris.
10. Sjakste T, **Paramonova N**, Lunins R, Limeza S, Sugoka O, Trapina I, Rumba-Rozenfelde I. Variability of the 14q proteasomal genes in populatuions and in the association with complex diseases. V International Meeting "Early events in Human Pathologies". Listvyanka, Baikal, Russia, Abstract book. 2012, 9-12 July, pp 33.
11. Limeza S; Rumba- Rozenfelde I; Sugoka O; **Paramonova N**; Trapina I; Sjakste T. "The heterozygosity at the PSMA3 c.543+138 locus is associated with obesity in children in Latvian population. Hormone research in paediatrics. 50th Annual Meeting of the European Society for Paediatric Endocrinology, Glasgow, UK, Abstracts. 2011, 25.- 28. septembris, 182 lpp.
12. **Paramonova N,** Kalendar R, Schulman A, Trapina I, Sjakste N, Sjakste T. Cloning of the human *PSMA6* gene promoter: haplotype analysis. Latvian University 69. Conference, Medicine section, Riga, Latvia, Abstract book. 2011, February 3, pp 64.
13. Trapiņa I. Sjakste T, **Paramonova N**, Lunins R, Sugoka O, Sjakste N. Genetic diversity of the 14q proteasomal genes in Latvian population. Latvian University 69. Conference, Medicine section, Riga, Latvia, Abstract book. 2011, February 3, pp 65.
14. Līmeža S, Rumba-Rozenfelde I, Rinkuža I, Sugoka O, **Paramonova N**, Trapina I, Sjakste T. Association between the *PSMA6* and *PSMA3* gene variation and obesity in children in Latvian population. Latvian University 69. Conference, Medicine section, Riga, Latvia, Abstract book. 2011, February 3, pp 66.

INTRODUCTION

Autoimmune and metabolic diseases cause numerous health and social problems throughout the world. Over the last decade increased interest of many researchers is directed to define the role of genetic sequence variation in the development of multifactorial diseases. Mutations in human genetic code can result in proteins with malformed proteins, altered functions, or missing proteins. Some of these changes may lead to disorders. Ubiquitin proteasomes system (UPS) is involved in removing of abnormal and misfolded intracellular proteins, and generation of antigenic peptides. The possible association of polymorphisms in proteasome subunits with human pathologies including metabolic and autoimmune diseases is intensively sought. We have focused our attention on juvenile idiopathic arthritic (JIA), bronchial asthma (BA) and obesity (OB). The candidate genes from proteasomes genes clusters located in the 14q11.2 - 14q23 chromosomal region, earlier investigated as diseases susceptible, was chosen for the current association study.

The cause of JIA is complex, involving both environmental and genetic risk factors. The latter could include structural variations in both human leukocyte antigens and non-HLA genes (Prahalad et al., 2008; Thompson et al., 2010). Due to the pleiotropic effect, which is a frequent phenomenon in human complex traits and diseases (Sivakumaran et al., 2011) some loci of susceptibility may be shared among many autoimmune and other immune-mediated diseases (Lee et al, 2014; Thompson et al, 2010). However, interplay of multiple risk alleles and/or genotypes and primary driver of these diseases remains still unclear.

Asthma is a chronic inflammatory disease caused by complex gene-gene and gene-environment interactions with hyper-responsiveness to various nonspecific stimuli (Broide, 2001; Galli et al., 2008) being to a large extent genetically heterogeneous in different human populations (Leung et al., 2014). The 14q11-24 chromosomal region was defined as one of well replicated asthma susceptibility loci (Hakonasron et al., 2002; Malerba et al., 2001; Mansur et al., 2004). The *PSMA6, PSMA3* and *PSMC6* proteasomal genes located in the region were implicated earlier in susceptibility to autoimmunity (Sjakste et al., 2004b; 2010), inflammation (Sjakste et al., 2007; Wang et al., 2013).

According to the recent data the ubiquitin-proteasome system is implicated in the pathogenesis of obesity. It has been demonstrated that in Southern Taiwan and Japanese population plasma ubiquitin and proteasome levels inversely correlated with a male BMI (Chang et al., 2009; Sakamoto et al., 2010). The mutation in the *PSMB8* gene has been reported to be associated with the autoinflammatory syndrome with lipodystrophy in Japanese

(Kitamura et al., 2011). Earlier the association of the *PSMA3* gene polymorphisms with susceptibility to obesity in Latvian was detected (Kupca et al., 2013).

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 only a few studies directly addressed to that objective: 14q13.2 microsatellite polymorphism in Latvian and Finland populations (Kalis et al., 2003; Sjakste et al., 2004b) and genotyping of the *PSMB5* gene single nucleotide polymorphism (SNP) in four American ethnic groups (Wang et al., 2008), for example.

Allele specific sequence functional motifs potentially could significantly affect particular gene expression, including UPS functionality in general and network of different genes and proteins those involved in ethnic specific adaptation to environment among the others. The functional consequences of the most number of mutations and polymorphisms in proteasome genes are not known. Several of them not associated with disease have been found in diseased tissues (Sjöblom, 2006). Understanding whether or not these proteasome mutations are important in disease development will require research to determine how these mutations affect proteasome function and how they affect the cell's physiology.

The present thesis is to investigated genetic diversity of proteasomes genes cluster located in 14q11.2 - 14q23 chromosomal region in population level and to reveal possible associations of investigated loci with autoimmune and metabolic diseases, namely juvenile idiopatic arthritis, bronchial asthma and obesity in Latvian population.

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.

The thesis is presented in a form of publication assortment. It covers investigation of subtype- and sex-specific association with JIA of genetic polymorphisms of proteasome gene cluster located in 14q11.2 - 14q23 chromosomal region: loci *PSMB5* (rs11543947), *PSMA6* (rs2277460 and rs1048990), *PSMC6* (rs2295826 and rs2295827) and *PSMA3* (rs2348071) (Publication II). Genetic diversity of above loci except the *PSMB5* (rs11543947) locus was investigated in Latvian, Lithuanian and Taiwanese population (Publication I) and on bronchial asthma association in Latvian and Taiwanese population (Publication III). *PSMB5* (rs11543947) and *PSMC6* (rs2295826 and rs2295827) loci in Latvians were genotyped on children obesity association (Publication IV).

Aim of the study

To investigate diversity of polymorphisms in proteasome gene cluster of human chromosome 14 in population level and their association with juvenile idiopathic arthritis, bronchial asthma and obesity in Latvian population.

Tasks of the study

1. To investigate and to compare distribution of *PSMA6 (rs2277460), PSMC6 (rs2295826 and rs2295827)* and *PSMA3 (rs2348071)* proteasomal gene variations in Latvian, Lithuanian and Taiwanese populations.
2. To compare susceptibility to bronchial asthma with genetic variations in the 14q11.2 - 14q23 proteasomal gene cluster in Latvians and Taiwanese in general and depending on sex.
3. To study subtype- and sex-specific association of the *PSMB5* (rs11543947), *PSMA6 (rs2277460), PSMC6 (rs2295826 and rs2295827)* and *PSMA3 (rs2348071)*14q proteasomal genes with JIA in Latvians.
4. To investigate the plasma proteasomes levels depending on JIA subtype- and sex.
5. To evaluate a possible association between genetic variations in the *PSMB5* and *PSMC6* genes and childhood obesity in the Latvian population.
6. To evaluate the eventual functionality of nucleotide substitutions and their potential to be involved in human genome evolution and geographical adaptation *in silico.*

1. MATERIALS AND METHODS

1.1. Description of the case/control groups participating in the study.

We have analysed seven collections of DNA samples. Three of them represented the control groups formed by samples representing Latvian, Lithuanian and Taiwanese populations. The case groups were presented by Latvian juvenile idiopathic arthritis, childhood obesity and bronchial asthma patients, and Taiwanese patients with bronchial asthma diagnoses. A more detailed characterization of the collections is given below.

Control groups:

Latvian control group.

LV population was represented by 191 (117 females) inhabitants who were hospitalized in various departments of "Bikernieki" Clinic at Riga Eastern Clinical University 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.

Lithuanian control group.

LT population was represented by 150 individuals (97 females) who 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.

Taiwanese control group.

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.

Patients groups (cases)

The blood samples from juvenile idiopathic arthritis and childhood obesity patients were collected in cooperation with prof. I. Rumba-Rozenfelde. These case collections were formed from 2003 onwards; the samples were collected in the P. Stradins Clinical University Hospital and Children Clinical University Hospital Clinic Gailezers in Riga, Latvia.

Latvian patient group with bronchial asthma.

One hundred two children (28 girls) under five represented Latvian asthma group. LV BA patients were enrolled from the outpatient clinic of P. Stradins Clinical University Hospital and Childrens Clinical University Hospital Clinic "Gailezers" in Riga, Latvia.

Taiwanese Bronchial asthma patient group.

One hundred fifty nine (69 girls) under three represented Taiwanese (TW) asthma group. TW study subjects were enrolled from elementary school for allergy diseases screen Taoyuan General Hospital, Taiwan.

All Latvian and Taiwanese patients were diagnosed with mild or moderate persistent asthma according to the guidelines of the Global Initiative for Asthma (GINA; http://www.ginasthma.org/local/uploads/files/GINA_Under5_Pocket_20091_1.pdf). Informed consent was obtained from all the study participants or their parents.

Latvian Juvenile Idiopathic arthritis patients.

Patients were 174 JIA children (108 girls) receiving consultation at the outpatient clinic of P. Stradins Clinical University Hospital and Children Clinical University Hospital Clinic Gailezers in Riga, Latvia. JIA was diagnosed and assignment of the JIA patients to subgroups was carried out according to the criteria of the International League of Association for Rheumatology (Petty et al., 2004). For association analysis both persistent and extended JIoA, and rheumatoid factor negative and rheumatoid factor-positive JIpA subgroups were combined in JIoA and JIpA groups of 107 and 55 patients, respectively. Twelve other patients were diagnosed as having systemic (n = 9), enthesitis-related (n =2), and psoriatic (n =1) arthritis.

Latvian Childhood Obesity patients group.

Patients were 94 obese children (1–17 years of age, 44% girls) who consulted an endocrinologist at the “Gailezers” Clinic of Children’s Clinical University Hospital in Riga, Latvia After excluding a comorbid disorders – diabetes mellitus, hypertension, obstructive sleep apnea, or psychiatric disease - or chronic use of medications that could be implicated in the aetiology of obesity, a diagnosis of obesity was established in accordance with WHO criteria, the body mass index (BMI in kg/m2), calculated and applied to age- and height-specific and gender-appropriate charts, was at least at the 97th percentile (de Onis M. et al., 2007). The study consisted of initial informed consent, followed by a questionnaire-based interview concerning family health history, a subsequent physical examination, and blood withdrawal.

Ethics:

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.

1.2. Analysed loci and genotyping methods

DNA extraction.

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 (1% agarose (Chempol, Czech Republic) and 1% gel in TAE buffer (pH = 8.5, 40 mM Tris, 20 mM acetic acid, 2 mM EDTA). Determining the concentration used BioPhotomter Eppendorf (Eppendorf, Germany). DNA samples were stored at −80 °C until use.

Measurement of plasma proteasome concentration (Publication II)

Plasma samples were available only for 23 JIA patients. These plasma samples were obtained from patients randomly chosen for plasma sampling during development of the DNA collection in the JIA study. Therefore, preliminary information on genotypes of these patients did not exist at the moment of the sampling. Blood was harvested on citrate anticoagulant, and plasma stored at -80°C. Plasma proteasome (p-proteasome) concentration was measured in triplicate per each sample using a standard 20S/26S Proteasome ELISA kit (BML-PW0575, ENZO Life Sciences) according to the manufacturer's protocols. Absorbance was read at 450 nm using a UV-Vis spectrometric plate reader. Results were expressed as concentration of proteasome protein in ng/ml determined by interpolation for the absorbance value using the generated 20S proteasome standard curve.

Marker choice

Due to limited data on the genetic diversity and susceptibility to diseases of proteasomal genes, several criteria were taken into account in choosing markers. These included the existence of previously reported findings on locus association with human health status, locus allele-specific potential to be functionally significant, locus variability in populations, Hardy-Weinberg expectations and others concerning mainly a genotyping technology. The rs2277460 and rs1048990 of the *PSMA6*, rs2295826 and rs2295827 of the *PSMC6* and rs23480071 of the *PSMA3* were previously studied on disease susceptibility (Alsmadi et al.,

2009; Banerjee et al., 2008; Barbieri et al., 2008; Bennett et al., 2008; Freilinger et al., 2009; Goncharov et al., 2009; Heckman et al., 2013; Ikeda et al., 2012; Liu et al., 2009; Ozaki et al., 2006; Sjakste et al., 2007; Wang et al., 2013; Kupca et al., 2013). The rs11543947 of the *PSMB5* gene was previously genotyped on genetic diversity only in HapMap populations. All loci fit all other mentioned criteria of marker choice. Table 1.2.1 summarizes information on the polymorphic loci genotyped in our study.

Table 1.2.1. Polymorphisms' description

Marker ID	Coordinates cytogenetic/genomic	Gene	Function	MAF/MA count	Association findings	
					Disease	Reference
rs11543947	14q11.2/ 14:23034812	*PSMB5*	Exon1: c.70C>T; Intron1: c.-112+300C>T	T=0.0359/ 180	OB	Publication IV
rs2277460	14q13.2:/ 14:35761573	*PSMA6*	Promoter: c.-110C>A (c.-109-1C>A)	A=0.0755/ 378	T2DM	Sjakste et al., 2007
					JIA,	Sjakste et al, 2009; Publication II;
					BA	Publication III
rs1048990	14q13.2/ 14:35761675	*PSMA6*	5'-UTR: c.-8C>G	G=0.1903/ 953	T2DM	Sjakste et al., 2007; Barbieri et al., 2008
					CVD,	Ozaki et al., 2006; Sjakste et al., 2007; Takashima et al., 2007; Bennet et al., 2008; Hinohara et al., 2009; Honcharov et al., 2009; Ikeda et al., 2012; Heckman et al., 2013; Wang et al., 2013
					JIA	Publication II;
					BA	Publication III;
					Cancer	Bachman et al., 2010
rs2295826	14q22.1/ 14:53174923	*PSMC6*	Intron 1: c.128-104A>G	G=0.1352/ 677	JIA,	Publication II;
					BA	Publication III;
rs2295827	14q22.1/ 14:53174981	*PSMC6*	Intron 1: c.128-46C>T	T=0.1078/ 540	JIA,	Publication II;
					BA	Publication III;
rs2348071	14q23/14: 58730626	*PSMA3*	Intron 7: c.543+138G>A; c.522+138G>A; c.576+138G>A	A=0.3482/ 1744	JIA,	Publication II;
					BA	Publication III;
					Obesity	Kupca et al., 2013

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

1.3. DNA genotyping.

The recommended nomenclature system (http://www.genomic.unimelb.edu.au/mdi/mutnomen/recs.html/) and 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. Genotyping methods and primer sequences used in LV and LT study are indicated in Table 1.3.1.

Table 1.3.1. Primers for DNA amplification and details of the genotyping technology.

Gene ID	Marker ID	Method / restrictase	Primer	Product size (bp) amplified / restricted
PSMB5	rs11543947	CAPS / *AasI*	F-5'- AAATGGTCTTTCGCATCTGG-3' R-5'- CTCGGCCAAGATTCATTGTT-3'	691 / 457 + 234
PSMA6	rs2277460	ASA	F-5'- ATGCAAGAGCGGAAGAAA[C]-3'	256
	rs1048990	CAPS / *RsaI*	F-5'- ATGCAAGAGCGGAAGAAA[A]-3' R-5'- TACCATGACAGGGCAATTCAG-3'	256 / 161 + 94
PSMC6	rs2295826	CAPS / *DdeI*	F-5'- GCTTAAACAAGTATTGCCGATCA-3' R-5'- AAGGAAGAAAATAAAAAGCATTACTT-3'	411 / 133 + 276+ 2*
	rs2295827			411 / 194 + 215 + 2*
PSMA3	rs2348071	CAPS / *TscAI*	F-5'- GTCTAAGGCAGGGATGTCCA-3' R-5'- ACCAGCTTTCCCATTCAGTG-3'	232 / 166 + 66

Abbreviations of the genotyping method: ASA - allele specific amplification; CAPS - cleaved amplified polymorphic site. Letters "F" and "R" given in the primer sequences, indicate forward and reversal primers respectively. Restriction enzymes are given in italic. Nucleotides genotyped by ASA are boxed.

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 sec, appropriate annealing temperature (55 – 61°C) for 45 sec, 72°C for 45 sec and a final extension at step at 72°C for 7 min.

DNA digestion by restriction enzymes was performed according to the producer 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 being in 61 bp distance between each other were genotyped simultaneously in one *DdeI* digestion reaction.

Each single locus genotype was represented by specific band pattern on agarose gel: the rs2295826 AA - 276 + 133 bp, AG - 411 + 276 + 133 bp, GG – 411; the rs2295827 CC – 215 + 194 bp, CT – 411 + 215 + 194 bp, TT – 411 bp giving easily distinguished specific band combinations for 2-locus genotypes: AA/CC – 215 + 133 + 61 bp; AA/CT – **276** + 215 + 133 + 61 bp; AA/TT – **276** + 133 bp; AG/CC – 215 + 194 + 133 + 61 bp; 2 x AG/CT - 411 + 215 + 133 + 61 bp or **276** + 215 + 194 + 133 bp; AG/TT – 411 + **276** + 133 bp; GG/CC – 215 + 194 bp; GG/CT – 411 + 215 + 194 bp; GG/TT – 411 bp. Marked in bold the 276 bp band and underlined genotypes were not observed in the study. (*) - fragment of 2 bp was not taken into account in above description of genotypes discrimination. Amplified and digested products were analyzed by electrophoresis in 1 – 3% agarose gel.

Taiwanese specimens were genotyped 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 (North Western European), HCB (Han Chinese) and JPT (Japanese) populations respectively and 5 locus genotypes were reconstructed for each individual (Publication I).

1.4. Data processing and statistical analysis

Documenting personalised genotyping data allowed determination of multi-locus genotypes, observed haplotypes, single locus genotypes (SLG) and allele frequencies were estimated by direct counting of genetic variants and deviations from Hardy–Weinberg equilibrium (HWE) were tested by ÷2 test (Rodriguez et al., 2009; http://www.oege.org/software/hwe-mr-calc.shtml). Inferred haplotypes prediction, haplotype sorting, estimation of the linkage disequilibrium and probability of recombination were

performed using the DnaSP software version 5.10.1 online tool at http://www.ub.es/dnasp. (Librado et al., 2009). Both the two-tailed Fisher's exact test and the χ^2 test were applied to evaluate the linkage between the rs2295826 and rs2295827 polymorphic sites at three *P*-value levels ($P < 0.05$; $P < 0.01$; $P < 0.001$). The Bonferroni correction included in the DnaSP analysis was taken into account to support the significance of the revealed disequilibrium ($\alpha' = 0.05$).

Deviation from the Hardy-Weinberg equilibrium and differences between cases and controls in allele, genotype and haplotype frequencies were evaluated by χ^2 and Cochran-Armitage trend test using XLSTAT 2013 software for Windows.

Genetic models for every individual locus were designed according to Lewis (Lewis et al., 2002). Contingency tables were 2 x 3 for the AA, AB, BB genotypes in general model; 2 x 2 for the AA, AB+BB and AA+AB, BB and AB, AA+BB genotypes in dominant, recessive and over dominant models respectively and A and B alleles in the multiplicative model where A is the major allele and B is the minor allele. Using an additive model, the AA, AB and BB genotype distribution was analysed using the Cochran-Armitage test for trend. An odds ratio (OR) more than 2 and less than 0.5 was considered to be clinically significant.

In case of genetic variations current study in population level (Publication I) DnaSP software version 5.10.1 was used to evaluates the nucleotide diversity, performs Tajima's D (Tajima et al., 1989) and Fu and Li's F* and D* (Fu et al., 1993) tests of neutrality, and evaluates the pairwise linkage disequilibrium (LD) between the loci (D and r2) and pairwise population differentiation (Fst). 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).

Stratification was performed by JIoA and JIpA ILAR subtypes and by sex (Publication II); by sex in bronchial asthmas association study (Publication III); by sex and familial subgroups in childhood obesity study (Publication IV).

Levels of p-proteasome were expressed as mean ± standard error of the mean (SE) for each sample to show the variability associated with the estimation, and as mean ± standard deviation (SD) to characterize the spread of a data set within the groups. Both SE and SD were calculated using the on line NCalculators (http://ncalculators.com/). Differences in p-proteasome levels between the groups were estimated by nonparametric Mann-Whitney and/or Kruskal-Wellis tests using XLSTAT 2013 software. *P* values of less than $5x10^{-2}$, $2x10^{-3}$ and $1x10^{-4}$ were considered to be of nominal, moderate and strong statistical significance, respectively (Manly et al., 2005).

1.5. SNPs functional analysis in silico (Publications I, II)

An eventual functional significance of the SNPs showing evidence of association was analysed *in silico* on sequence similarity to transcription factors binding sites (TFBSs) using Genomatix software, MatInspector, Release 7.4 online tool, at www.genomatix.de (Cartharius et al., 2005). Only parameters with core/matrix similarity of more than 1.000/0.800 were taken into account. Splicing signals were predicted by Human Splicing Finder (HSF) Version 2.4 (http://www.umd.be/HSF; Desmet FO et al., 2011) with standard threshold values for branch point, donor and acceptor splice sites, enhancer, silencer, hnRNP and other splicing motifs. Sequence similarity to mature microRNAs and hairpin precursors was evaluated, and microRNA targets prediction was done using miRBase (http://www.mirbase.org/index.shtml; Kozomara et al., 2011) and miRNAMap (http://mirnamap.mbc.nctu.edu.tw/index.php; Hsu et al., 2008) online tools, respectively.

2. LITERATURE REVIEW

2.1. Types of polymorphisms in human genome and their importance for disease research.

Over the last decade, significant improvements have been made in genotyping efficiency, sequencing technology, and statistical methodology, providing researchers with better opportunities to define the role of sequence variation in the development of human diseases (Li. et al., 2013; Sonn et al., 2013; Wang et al., 2013). Many human diseases are now known to have a genetic component. All humans start their lives with germ-line mutations inherited from their parents. However, the human genetic code is constantly subjected to mutations which can happen during cell division or after exposure to environmental factors such as UV radiation, chemicals, or viruses. These mutations can result in proteins with altered functions, malformed proteins, or even missing proteins. Some of these changes that occur due to a particular mutation have no effect on biological function, some may be beneficial, and some may lead to disease.

2.1.1 Single-nucleotide polymorphisms (SNPs)

Polymorphisms occur more often (frequency of 1% or greater) in the general population than mutations (Wood et al., 2013; Abecasis et al., 2013). Single-nucleotide polymorphisms (SNPs) are the most frequent type of variation in the human genome, account for 90% of human DNA polymorphisms (Gomes et al., 2013) and they provide powerful tools for a variety of medical genetic studies. (Wang et al., 1998)

Most SNPs have two alleles which are named “major” or “common” and “minor” or “rare” based on their observed frequency in the general population. Several genotypes are possible at each SNP, because chromosomes are both maternal and paternal in origin: homozygous for the common allele, heterozygous, or homozygous for the rare allele. More than 10 million SNPs occur in the genome, one for every 100~300 bases (HapMap Consortium, Nature, 2003). Whole genome sequencing of large sample numbers is not always effective for finding SNPs, as many rare SNPs occur only once (“singletons”) or twice (“doubletons”) in the analysed samples The haplotype refers to an individual collection of short tandem repeat allele mutations at adjacent locations (loci) that are inherited together (Gomes et al, 2013).

Smaller research labs are now able to analyse multiple genes belonging to the same pathway instead of analysing a single polymorphism on a single gene. These advances have led to the discovery of new polymorphisms on proteasome genes that are linked to major

human diseases. SNPs can be used as genetic markers for identifying susceptible with disease genes by linkage studies in families, linkage disequilibrium, and association analysis of case/control study. Most investigations use dense maps of SNPs as well as the haplotypes derived from these polymorphisms. Determining the causal relationship between SNP or haplotype and disease is currently a major researching theme. These genetic variations are important for genetic diversity within the population.

Genome scan approaches to identify regions associated with diseases are now much more efficient (HapMap Consortium, Nature, 2003) The HapMap base contains SNPs data and maps of SNPs haplotype blocks, allowing users to select a group of SNPs to investigate a possible association between known genomic regions and the studied diseases.

So, previously were identified four non-HLA type 1 diabetes risk loci: *INS, CTLA4, PTPN22*, and *IL2RA1-4* resided in the *HLA* region on chromosome 6p21. (Jeffrey et al., 2009)

Recently, the application of genome-wide SNP typing technology to large sample sets and comparisons with results from other immune-mediated diseases have provided convincing support for 19 additional type 1 diabetes loci 5-13, all with allelic odds ratios (OR's) of less than 1.3. (Jeffrey et al., 2009)

Genome wide association studies have identified the susceptibility locus for lung cancer to be present on chromosome 15. This locus 15q24-15q25.1 contains not only the genes of nicotinic acetylcholine receptor subunits (*CHRNA3* and *CHRNA5*), which could be good candidate genes for lung cancer, but also *PSMA4*, the proteasome alpha type subunit isoform 1 gene. This region was found to be strongly associated with the risk of lung cancer in the Caucasian and African American populations (Amos et al., 2010; Hansen et al., 2010; Liu et al., 2008).

The *IFIH1* (interferon induced with helicase C domain 1, also known as MDA5, or melanoma differentiation-associated gene 5) linkage disequilibrium block on chromosome 2q has also been found to be associated with T1DM in GWAS and increased gene expression is associated with risk of T1DM (Nokoff et al., 2013). Meta-analysis of GWAS results revealed an association of end-stage renal disease (ESRD) with the rs7583877 in the RNA-binding protein *AFF3* gene and a chromosome 15q26 intergenic SNP locating between two growth and development-related *RGMA* (repulsive guidance molecule a) and *MCTP2* (multiple C2-domains with two transmembrane regions 2) genes. Functional data suggest that *AFF3* influences renal tubule fibrosis via the transforming growth factor-beta (TGF-β1) pathway. A whole genome linkage scan by microsatellite markers provided significant evidence of linkage between the T1DM and multiple sclerosis (Pitzalis et al., 2008). It should be

mentioned that some genetic predisposition factors apparently are not spotted in GWAS studies, although sufficient association is revealed in studies with individual genes.

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. GWAS studies have identified alleles related to complex disorders; however some of these alleles seem to be associated with the disease only in certain populations. Some polymorphisms were found in strong association with diseases in one or a few ethnic groups but not confirmed by the results in other ethnic groups. So, SNP rs1048990 *(PSMA6)* proteasomal genes was found in association with decreased risk of ischemic stroke in both Caucasians and African Americans (Heckman et al., 2013). No association between the *PSMA6* polymorphism and CAD was found in the North Indians (Banerjee et al, 2009). Case/control study from Japanese and Korean populations for *PSMA6* genotypes showed no evidence of the association with CAD (Hinohara et al., 2009).

2.1.2 Microsatellites (MSs)

MSs are repeating sequences of 2–6 base pairs of DNA (King et al., 2012). They are used as molecular markers in genetics for kinship, population, and other studies. They can also be used for the studies of gene duplication or deletion, marker-assisted selection, and fingerprinting. Microsatellites are distributed throughout the genome (King et al., 2012). Being variable genetic elements, microsatellites provide a potent tool for the individual characterization of genomes. Variability is generated due to replication slippage caused by mismatches between DNA strands while being replicated during meiosis (Tautz et al., 1994), and the event can occur once per 1000 generations (Weber et al., 1993). This slippage is much more common compared with point mutations (Jarne et al., 1996).

Microsatellite repeats mutagenize human genomes and alter the human genomic landscape across generations (Grandi et al., 2013). The utility of microsatellites has been demonstrated by the study comprising 2058 germline changes discovered by analysing 85 289 Icelanders at 2477 microsatellites. The paternal-to-maternal mutation rate ratio is 3.3, and the rate in fathers doubles from the age 20 to 58, whereas there is no association with age in mothers. Longer microsatellite alleles are more mutagenic and tend to decrease in length, whereas the opposite is seen for shorter alleles (Sun et al., 2012).

Microsatellites remain highly informative and useful measures of genomic variation for linkage and association studies despite the fact that general preference is given to single-nucleotide polymorphisms (SNPs). Microsatellites are much more genetically diverse

compared to SNPs; they generate greater haplotype diversity (Gulcher et al., 2012). Although mostly used as structural genetic markers, microsatellites perform several functions in the genome, which are still far from being completely understood.

However, actually it has become clear that MSs and their flanking regions are involved in multiple gene and genome functions. MSs are known to form nuclear matrix anchorage sites (Boulikas et al., 1993), tissue-specific matrix attachment sites (Lenartowski et al., 2002), and binding sites with vimentin and glial fibrillary acidic protein (Tolstonog et al., 2005) and give rise to complex DNA spatial structures of extreme functional significance (Tolstonog et al., 2005; Li et al., 2003; Rothenburg et al., 2001). MSs appear to be important components of insulators (Filippova et al., 2001), silencers (Rothenburg et al., 2001), and enhancers (Bassuny et al., 2003). MSs are also involved in the regulation of alternative splicing (Hui et al., 2005), mRNA stability (Lee et al., 2004), and recombination and repair (Wang et al., 2007 Chin et al., 2007). Expansions of microsatellite DNA repeats cause nearly 30 developmental and neurological inherited disorders (Kim et al., 2013).

2.2. Ubiquitin proteasome system and UPS-related diseases

Recent advancements in cell biology show that the function of the cells equally important is the proper protein biosynthesis, as well as their efficient degradation. In eukaryotes, processing and degradation of vast majority of regulatory proteins are mediated by ubiquitin-proteasome system (UPS).

2.2.1. The proteasome

The proteasome is a multicatalytic enzyme which is highly conserved and ubiquitous. There are found in archaebacteria as well as the nucleus and cytoplasm, near the endoplasmic reticulum and even in the centrosome of all eukaryotic organisms (Fabunmi et al., 2000).

Proteasomes, key UPS enzymatic complex possess several types of peptidase, endoribonuclease, protein-chaperone and DNA-helicase activities (Konstantinova et al., 2008; Wang et al., 2007; Yano et al., 2005) allowing to control strictly and to coordinate all steps of gene expression, genes' and proteins' networks and processes of genome-environment interaction. Proteasome is involved in the cellular stress response mechanisms, timely regulatory factors degradation provides the cell response to DNA damage, hypoxia and heat shock, deviations from the correct course of this process may be seen as pathological (Flick et al., 2012; Wang et al., 2006). Proteasomal degradation removes denatured, misfolded, damaged or improperly translated proteins from cells and as well as regulating the level of proteins such as cyclins and transcription factors (Nandi et al., 2006; Hershko et al., 1998).

The removal of any proteasome gene is lethal in eukaryotes (Hamazaki et al., 2007; Bedford et al., 2008). The importance of the proteasome in cellular functions is exemplified by experimental evidence which suggests that the proteolytic capacity of the proteasome in certain tissues declines with age and that this decline in proteasome activity is related to the lifespan of the organism (Carrard et al., 2002; Vernace et al., 2007; Dasuri et al., 2009; Chondrogianni et al., 2000; P´erez et al., 2009); Aging cells have increased levels of damaged proteins, possibly increasing the load on the proteasome (Hipkiss et al., 2006).

2.2.2 Proteasomes components and structure

Proteasome is based on a 15x10 nm cylinder with a molecular weight of approximately 700kDa. Subunit orientation and detailed structure of the different subunits of the 20S catalytic particle was resolved 18 years ago (Groll et al., 1997). It consists of four rings, each of which consists of the seven subunits with a molecular weight of 20 to 36 kDa. The proteasome from the archeon *Thermoplasma acidophilum* is a prototype for the quaternary structure and topology of the enzyme (George et al., 1999). The 2 identical outer rings are composed of the α-type subunits and the 2 identical inner rings, of the β-type subunits. Therefore, the general structure of the proteasome complex is $\alpha_{1\text{-}7}\beta_{1\text{-}7}\beta_{1\text{-}7}\alpha_{1\text{-}7.}$ The α-rings together with the 19S regulatory particles form a narrow channel through which only denatured proteins may pass. The β-rings, each of which contains 3 active sites, form a proteolytic chamber where proteins are degraded in a progressive manner.

The proteolytic sites of the β-rings differ in their substrate specificity and proteolytic activity; therefore, they are named after enzymes that show similar activity or specificity: a “chymotryptic-like” activity with preference for tyrosine or phenylalanine at the P1 position; a “tryptic-like”activity with preference for arginine or lysine at the P1 position; and a “postglutamyl” hydrolysing activity with a preference for glutamate or other acidic residues at the P1 position (Orlowski, et al., 1990; Coux , et al., 1996).

During protein degradation in the proteolytic chamber, peptides of 3-25 amino acids in length are generated. The above-described complex forms the 20S Core Particle (CP) - 20 S proteasome (S - Svedberg coefficient characterizing the particle sedimentation rate). (Naujokat et al., 2002; Kisselev et al., 2003; Ciechanover et al., 2012). Polyubiquitinated proteins cannot be degraded directly by the 20 S proteasome. After attachment of at least four ubiquitin molecules the target protein is recognized by specific sub-units of 19S regulatory particle (RP) of the proteasome. 19S RP known alternatively as PA700 (Ma et al., 1994), ball (Hoffman et al., 1992), 19 S cap (Glickman et al., 1998) or μ-particle (Udvardy et al., 1993) is a multi-protein structure that caps the two sides of the core particle (CP) of the proteasome.

19S is made of two sub-complexes called the lid and the base and a total of 17 peptide molecules. Six of them possess ATPase activity while the 11 others are non-ATPases. At 20 S proteasome joins 19 S ATPase complexes that recognize the substrate. These both complexes are formed of 26 S proteasome representing a complex of 2.5 MDa (Gomes et al., 2006; Tomko Jr. et al., 2013; Kish-Trier et al., 2013).

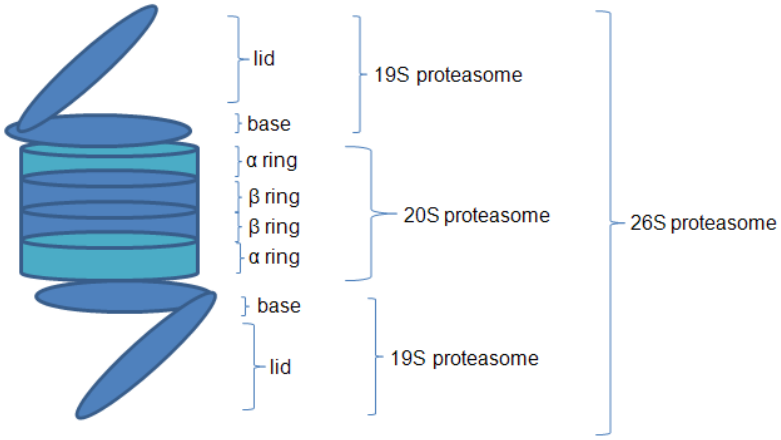


Figure 2.2.2.1. Schematic diagram of the 26S proteasome (I. A. Voutsadakis, 2010)

Intracellularly, multiple forms of the proteasome with different combinations of activators co-exist (Figure 2.2.2.2).

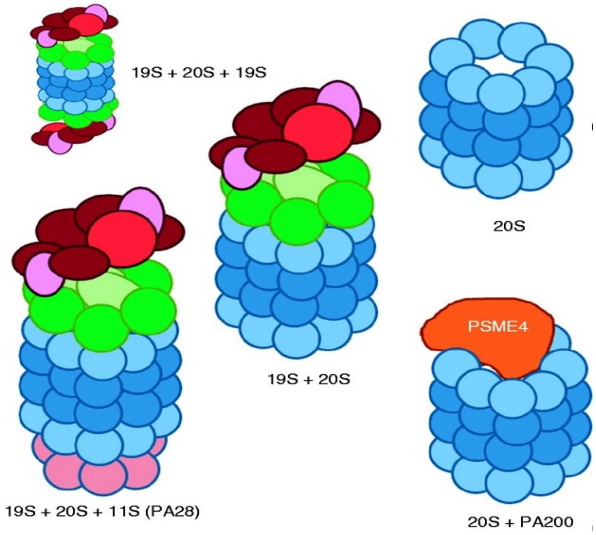


Figure 2.2.2.2 Schematic diagram of different forms of the proteasome. (http://www.reproduction-online.org/content/142/1/1/F1.large.jpg)

The 26S proteasome can exist with one or two 19S caps, immunoproteasomes containing one or two 11S caps, proteasomes containing the 20S proteasome with one or two PA200 caps (in the nucleus only), and hybrid proteasomes which contain different combinations of 20S and activators (Sutovsky et al, 2011; Gomes V et al, 2013). In the immunoproteasome the 19S regulator is replaced by an alternative 11S regulator (*PSME*), composed of *PSME1,*

PSME2 and *PSME3* subunits previously indicated as *PA28* alpha, beta and gamma respectively. Three *PSME1* and three *PSME2* subunits combine to form a heterohexameric ring. Several others ATP-independent regulatory complexes of different functions and protein composition have been recently discovered and characterized (Sorokin et al., 2009).

2.2.3 The Proteasome Proteolytic Pathway

The proteasome becomes capable of binding the substrate after the ubiquitination process. The ubiquitin-proteasome proteolytic pathway starts from ubiquitin, as proteasomes cannot recognize a protein that is not targeted. Degradation of a protein via the ubiquitin-proteasome system (UPS) involves 2 successive steps (Ciechanover et al., 2000):

1) covalent attachment of multiple ubiquitin molecules to the substrate;

2) degradation of the tagged protein by the 26S proteasome and recycling of ubiquitin. Ubiquitin is an 8 kDa protein of 76 amino-acids that has taken its name from its ubiquitous presence in cells. Its covalent link to a target protein is a signal for different fates for this target protein. Ubiquitin binds to the target protein with isopeptide site between the ubiquitin carboxyl terminal glycine and internal lysine of target substrate. Ubiquitin molecules, with additional links successively connected to one another to create a poly ubiquitin chain which serves as a recognition signal of the proteasome-ubiquitin in proteasome system. Modification of the target protein with ubiquitin or ubiquitin-like protein changes the target protein surface that reduces protein stability, activity and association with other proteins and its localization in the cell (Ciechanover et al., 2012; Bendotti et al., 2012).

Conjugation of ubiquitin to the protein substrate proceeds via a 3-step cascade mechanism carried out by enzymes E1, E2 and E3 (Fig 2.2.3.1).

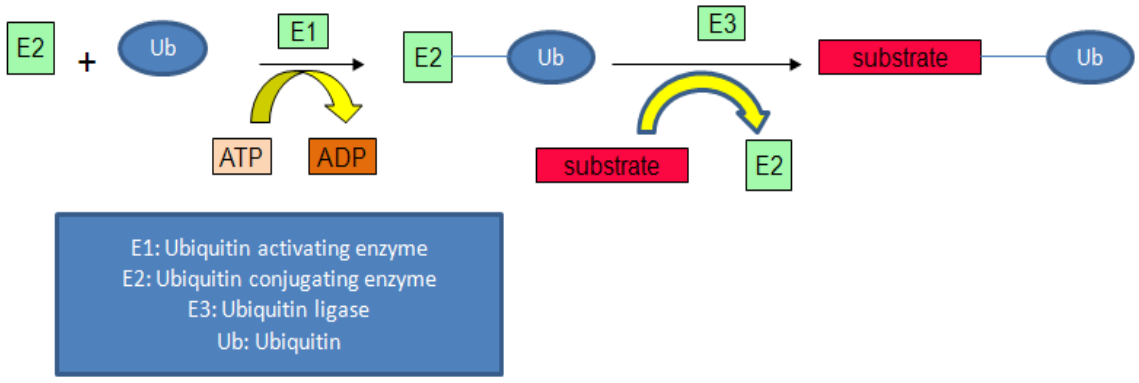


Figure 2.2.3.1. Enzymatic cascade of ubiquitination. (I. A. Voutsadakis, 2010)

The process of ubiquitination begins with adenosine triphosphate ubiquitin-dependent activation by the enzyme E1 (ubiquitin activating enzyme). Ubiquitin is linked to cysteine

residue E1, forming E1 -S ~ ubiquitin and then transferred to one or more forms or E2 ubiquitin enzymes. Third, the addition of ubiquitin to the substrate protein is catalysed by one of several E3 - ubiquitin ligases, which is a large group of proteins. After degradation deubiquitinylated enzyme separates the ubiquitin residues from substrate and ubiquitin can be reused (Ciechanover et al., 2012). The structure of the ubiquitin conjugating and activating system is hierarchical. A single E1 protein activates ubiquitin required for all modifications. It can transfer ubiquitin to several species of the E2 enzymes. More than 20 genes of E2 are discovered. Each E2 can transfer ubiquitin to one or several E3 proteins. E3 enzymes may belong to three different classes: proteins RING (Really Interesting New Generation proteins), U-box proteins and HECT-domain proteins. There are about 1000 E3-encoding genes that have been identified in the human genome. E3 proteins play a very important role in the ubiquitin-conjugating system, because they act as substrate-binding elements, which are responsible for specificity and selectivity of the ubiquitin-proteasome system (Glickman et al., 2002, Schwartz et al., 2009). Conjugation of ubiquitin may also be responsible for regulation of protein localization, functions and their interactions. Furthermore, ubiquitination plays an important role in modification of histones (Anna et al., 2006).

2.2.4. Monogenic diseases associated with UPS defects.

„E3 diseases"

E3 protein diversity determines their role in pathologies. Malfunction of E1 or E2 enzymes should be fatal for organisms, on the contrary silencing of one E3 can be partially compensated by analogous enzymes, and however such a deficiency will not remain without consequences, it causes the pathology. A distinct number of pathologies are caused by deficiency of some E3 enzyme, we can denote then as "E3 diseases."

Currently the role of E3 defects is better understood in neurology. About 17 different E3 enzymes are involved in development of nervous system. The signal transfer functions are regulated by 11 enzymes. (Baptista et al., 2012). Angelman syndrome provides a good example of the case when one of the enzymes is not active. This inherited syndrome manifests in patients as mental retardation, speech difficulties, unsteady gait, they often clap hands, laugh, and look happy. Patients inherit a 4 Mb large (4 million base pairs) 15q11-13 deletion in Chromoosme 15 from the mother; the Ube3a gene is lost among the other genes. This gene encodes E6-AP ubiquitin ligase, normally expressed in the cerebellum and hippocampus. Similar symptoms develop when Ube3a gene mutation is formed. A number of E6-AP substrates has been identified, but the mechanism of pathogenesis is not well understood yet (Buiting et al., 2010; Mabb et al., 2011; Williams et al., 2010; Baptista et al., 2012).

Williams-Beuren syndrome (WBS) is caused by a Chromosome 7q11.23 deletion, determining a loss of 25 genes, including the E3 ligase TRIM50 (Micale et al., 2008). The disease is characterized by many developmental and neurological defects. E3 enzymes CUl4B HUWE1 and is also involved in a number of undefined mental retardation cases (Baptista et al., 2012). E3 enzyme UBE3A, PARK2, RFWD2, FBXO40 deficiency may be an etiological factor for autism (Baptista et al., 2012). Recently described mutation of the *UBR1* E3 enzyme gene causes the typical Johnson-Blizzard syndrome or ectodermal dysplasia: abnormal development of the pancreas, nose and skull dysplasia, deafness and mental retardation (Fallahi et al., 2011). But overexpression of the TTC3 E3 ligase encoded in the "critical zone" for Down's syndrome can be responsible in part for neurological symptoms of this syndrome (Berto et al., 2007). One of the early forms of Parkinson's disease develops due to mutations in *PARK2 E3* ligase gene (Baptista et al., 2012). Sporadic Parkinson's disease patients have a reduced expression of E3 ligases component in Skp1 protein complex. The increase of this protein expression could contribute to the increase of disease treatment efficiency (Mandel et al., 2012).

Two hereditary cancer syndromes could be also classified as "E3 diseases". Von Hippel Lindau syndrome is characterized by E3 ligase VHL gene mutation that results in not-degradation of hypoxia induced transcription factor in normoxia conditions, this leads to hyperplasia and malignization in several organs (Bader et al., 2012). ACP tumour suppressor, which is also a E3 enzyme is responsible for the development of familial adenomatous polyposis. (Benanti et al., 2012).

"Deubiquitination diseases"

This group is formed by Machado-Joseph Disease, where incomplete deubiquitination of ATXN3 peptide accumulation causes neurotoxic effects (Costa et al., 2012). USP14 ubiquitin-specific protease defect causes ataxia in mice (Wilson et al., 2002). Human enzyme with a similar function (UCH-L1 or PARK5) is absent in some patients with sporadic forms of the Parkinson's disease (Belin et al., 2008).

"Immunproteasomes" and their defects.

Induction of inflammatory cytokines such as interferon-γ or tumour necrosis factor α (TNF-α) stimulate the "immunoproteasome" formation. Replacements of three β-subunit proteins take place. Proteasome proteins PSMB1, PSMB2 and PSMB5 are replaced by PSMB9, PSMB8 and PSMB10 proteins. The "immunoproteasome" participates in the class I major histocompatibility complex protein formation and antigen presentation. Three formerly unknown monogenic diseases caused by mutations in the *PSMB8* gene were described recently. These are Nakajo-Nishimura syndrome (Kanazawa et al., 2012), this syndrome is characterized by frostbite-like bluish erythematous patches on the skin and progressive partial lipodystrophy, JMP syndrome (joint contractures, muscle atrophy, microcytic anaemia, and panniculitis-induced lipodystrophy) - joint contractures, muscle atrophy, subcutaneous fat rounds of inflammation-induced lipodystrophy (Agarwal et al., 2010) and CANDLE syndrome (chronic neutrophil dermatosis with atypical lipodystrophy and elevated temperature) - a chronic atypical neutrophilic dermatosis with lipodystrophy and elevated temperature (Goldbach-Mansky et al., 2012). Lipodistrophy is a common feature of the three syndromes, other symptoms vary. However, all the three diseases are caused by c.224C> T mutation in exon 2 of the gene *PSMB8,* the mutation results in Thr75Met replacement in the protein. In some CANDLE syndrome patients the c.405C> A mutation identified in the same gene caused formation of truncated peptides.

"Bad" proteasome substrate pathologies

In a group of diseases inability of proteasomes to digest some substrate is a common step of pathogenesis. First of all these are polyglutamine tract diseases. Polyglutamine is found in a large number of proteins, it should have certain physiological functions. These proteins are degraded by ubiquitin-proteasome system, but in the course of aging, exposure to toxins, and in some other cases, the proteasome is no longer able to degrade aggregates formed bylong polyglutamine tract (Chen et al., 2005a, 2005b). Diseases can be modelled by treating the cells with silicon nanoparticles (Chen et al., 2005a). About ten polyglutamine tract diseases have been described, Huntington's disease is the most studied disease of this group; in this

case the neuronal protein huntingtin accumulates in neurons. In Kennedy's disease polyglutamine tract of androgen receptor provokes development of pathology (Li et al., 2011; Rusmini et al., 2010). Proteins with polyglutamine tract are degraded not only in the UPS sustem, but also in process of autophagy, stimulation of this process can have a positive impact on the disease symptoms (Jimenez-Sanchez et al., 2012). Liddl syndrome was described as a state of hypertension, hypokalaemia and metabolic alkalosis with low aldosterone and renin levels. It turned out that Liddl's syndrome is caused by a mutation in renal sodium channel (*ENAC*) beta subunit gene. The E3 ubiquitin ligase Nedd4 cannot recognize this mutated channel protein; it is not degraded in the ubiquitin-proteasome system. Excessive sodium reabsorption and water retention happen as result (Rotin et al., 2008).

Some forms of Alzheimer's disease can be also included in the "bad proteasome substrate disease" definition. Amyotrophic lateral sclerosis, a rare disease which begins around the age of 60 years and manifests itself ad progressive paralysis, muscle degeneration and disability, also develops due to the overload of ubiquitin-proteasome system. The symptoms develop due degeneration of motor neurons. Patients have a mutation in superoxide dismutase (*SOD1*) gene, the enzyme eliminates free radicals. Proteasome cannot cope with the mutant protein processing, resulting in its aggregation, "immunproteasomes" are over expressed in the same time (Bendotti et al., 2012). "Poor substrate" components can modify the severity of lysosomal diseases. Gaucher disease is caused by mutations in the beta- glucocerebrosidases gene, resulting in accumulation of glycol ceramides in lysosomes. It turns out that some mutations of this gene slightly alter the protein structure, which makes it difficult to degrade (Bendikov-Bar et al., 2012), more severe symptoms develop as result.

Autoimmune disorders

Over the past ten years the hypothesis about the ubiquitin-proteasome system decisive role in antigen presentation and immune response has been confirmed many times (Kloetzel et al., 2001). A new special form of the proteasome - thymoproteasomes found exclusively in the thymus cortex, and probably involved in T cell selection was recently described (Sutoh et al., 2012; Tomaru et al., 2009). This indicates an even greater role of proteasomes in the immune response and justifies the study of proteasome system activity in patients with autoimmune diseases. In patients with rheumatoid arthritis (Egerer et al., 2002) and psoriatic arthritis (Henry et al., 2011) concentration of proteasome in blood plasma is increased. Antibodies against the 20S proteasome were found in patients with Sjögren's syndrome (Feist et al., 1999), lupus erythematosus (Arribas et al., 1991; Colmegna et al., 2008; Feist et al., 1996), multiple sclerosis, psoriatic arthritis (Colmegna et al., 2008), autoimmune

cardiomyopathy (Voigt et al.; 2010), autoimmune myositis (Feist et al., 1996). Ubiquitin-proteasome system is also involved in other auto-antigen and nuclear protein processing. Proteasome degrades many nuclear proteins, including histones (Dino Rockel, 2002). Toxic substances in the nucleosome core lead to accumulation of abnormal degradation of proteasome substrates; these proteins acquire auto-antigenic properties. (Chen et al., 2002; 2005). Proteasome inhibitors are used for treatment of rheumatoid arthritis and other autoimmune diseases (Chitra et al., 2012).

Neurodegenerative diseases

Abnormal activity of the ubiquitin-proteasome system may be involved in the pathogenesis of Alzheimer's disease; however this pathway is not understood yet in detail. In neurons of Alzheimer's patients mainly Aβ un τ proteins resistant to proteasomal degradation are accumulated. Reduced proteasomal activity in neurons of patients is also reported, mutated ubiquitin is found sometimes. Lysosomal activity in neurons of Alzheimer diseases patients is increased in the initial stages of the disease; the late stages are characterized by increased autophagy process (Ihara et al., 2012; Dennissen et al., 2012). Mutations in ubiquitin gene were found in patients with amyotrophic lateral sclerosis. Huntington's disease is a classic "bad substrate" disease example; however diverse defects of the ubiquitin-proteasome system are characteristic of different forms of Parkinson's disease (Dennissen et al., 2012).

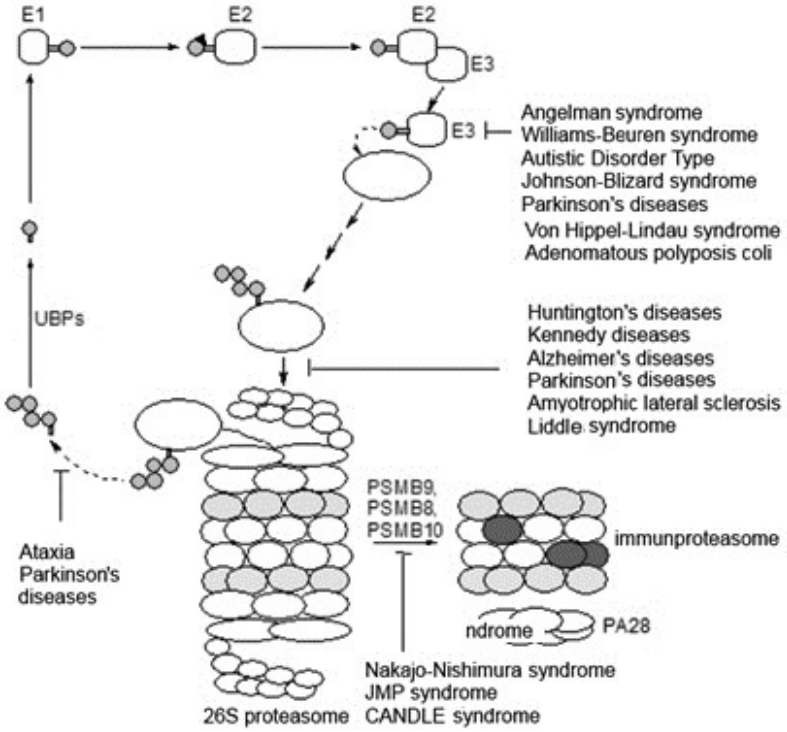


Figure 2.2.4.1. The potential proteasome malfunctions in pathology (Sjakste et al., 2013).

2.3. UPS system genes and their polymorphisms.

In the human genome, there are many genes that encode the proteins of the ubiquitin-proteasome system. In addition, the 26S proteasome is composed of proteins encoded by at least 34 independent genes (Nandi et. al., 2006). Distribution of the proteasomal genes over the human genome displays the tendency to their clustering in some chromosomes including the long arm of chromosome 14 where 9 related loci have been localized including the structural genes encoding two of beta (*PSMB5* and *PSMB11*) and of alpha (*PSMA3* and *PSMA6*) subunits of core 20S proteasome, ATPases (*PSMC1* and *PSMC6*), 11S non-ATPase activators (*PSME1* and *PSME2*) and the *PSMA3P* pseudogene.

Possible links between proteasomal gene polymorphism and human pathologies were first reported in 1994 (Van Endert et. al., 1994) and association between *LMP2* (*PSMB9*) gene, which encode beta type 9 subunit of proteasome, and juvenile rheumatoid arthritis was reported in 1995 (Pryhuber et. al., 1996). In 2006, Ozaki et al. published evidence of a possible association between *PSMA6* gene polymorphism and myocardial infarction (Ozaki et. al., 2006). After this study many other studies were performed in different populations, and study of the proteasomal gene polymorphisms has become popular among scientists. The study in a Chinese population also reported that the *PSMA6* variant rs1048990 could be a risk factor of MI and confirmed the association between polymorphism of the *PSMA6* gene and the development of myocardial infarction (Liu et. al., 2009). However, the similar studies in other populations showed controversial results. Bennet et al. concluded that the risk of rs1048990 polymorphism for myocardial infarction was unlikely to be great in Western populations (Bennett et. al., 2008). Moreover, none of the single nucleotide polymorphisms of the *PSMA6* gene was found to be associated with coronary artery disease in a study carried out in Saudi Arabia (Alsmadi et al., 2009;). However, a borderline association of the rs1048990 with diastolic blood pressure was detected in United Kingdom population (Bennett et. al., 2008).

"Immunproteasome" PSMB8 protein and gene *PSMB9* polymorphism, were the best studied in relationship with autoimmune diseases. The 14q11, 14q13, and 14q22-32 regions harbouring the listed loci, have been previously reported as susceptible to autoimmune (Arya et al. 2009; Burgner et al. 2009; Chistiakov et al. 2004; Cornelis et al. 1998; Sjakste et al. 2004b; Stuart et al. 2010; Tomer et al. 1997) and other multiple diseases (Banerjee et al. 2008; Barbieri et al. 2008; Bennett et al. 2008; Wang et al. 2013). Some of these proteasomes genes showed association with type 2 diabetes (Sjakste et al., 2007), Graves' disease (Sjakste et al., 2004a) as well as myocardial infarction and blood pressure (Sjakste et al., 2007). Proven association between *PSMA6* gene microsatellite allele HSMS006 $(TG)_{22}$ and type 2 diabetes

(Sjakste et al., 2007, 2008), multiple microsatellite alleles were found in associating with Graves' disease (Sjakste et al., 2004a) and juvenile idiopathic arthritis (Sjakste et al. , 2010). Another SNP is associated with juvenile idiopathic arthritis (Trapiņa et al. 2009). In recent years appeared a new investigating related to *PSMA6* gene and located near a gene polymorphisms, which were found in association with diseases (Ozaki et al. 2006; Kim et al. 2008; Stuart et al. 2010; Bachman et al. 2010; Barbieri et al. 2008; Liu et al. 2012; Goncharov et al. 2009; Liu et al. 2009; Ozaki et al. 2006; Wang et al. 2013; Burgner et al., 2009; Thompson et al., 2010; Arya et al. 2009). Our research shows that juvenile arthritis is associated with proteasomes *PSMA6* gene stationed near the *KIAA0391* genes genotype $(AC)_5$ $AT(AC)_{17}$ / $(AC)_5$ $AT(AC)_{19}$ and genotypes $(CA)_{14}$ / $(CA)_{19}$. In addition, our data point to a single genotype - *PSMA6* genotype $(TG)_{18}$ / $(TG)_{21}$ and $(TG)_{22}$ / $(TG)_{23}$ potential role in the disease in certain individuals (Sjakste et al., 2010). Promising results were obtained in studies in type 1 diabetes, multiple sclerosis, asthma and obesity.

2.3.1 Proteasomal genes genetic diversity on population level

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., 2004b) 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 2.3.1.1. for references).

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., 2004b, Trapina et al., 2009), British (Bennett et al., 2008; Freilinger et al., 2009), 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.

Table 2.3.1.1. Genotype and allele frequencies of the rs1048990 in different ethnic groups over the world

Population	MAF (G)	Genotype frequency			No. of controls	Study or project	Reference
		CC	CG	GG			
YRI	0.017	0.967	0.033	-	60	HapMap	NCBI resources
Latvian	0.089	0.827	0.168	0.005	191	Population	Publication I
	0.148	0.732	0.241	0.027	118	JIA	Trapina et al., 2009
	0.086	0.829	0.171	0	105	DMT2	Sjakste et al., 2007
	0.107	0.796	0.193	0.014	414		Mean for Latvians
CEU	0.117	0.783	0.200	0.017	60	HapMap	NCBI resources
Ukrainian	0.133	0.762	0.211	0.027	208	EH	Honcharov et al., 2009
South Italian	0.149	0.748	0.205↓	0.046↑	151#	MI &DMT2	Barbieri et al., 2008
Lithuanian	0.153	0.707	0.280	0.013	150	Population	Publication I
British	0.149	0.730	0.243	0.027	918	IS	Freilinger et al., 2009
	0.156	0.708	0.271	0.021	2713	MI	Bennet et al., 2008
	0.171	0.691	0.282	0.027	282	MI	Sjakste et al., 2007
	0.159	0.710	0.253	0.025	4195		Mean for British
Germany	0.169	0.683	0.297	0.021	725	IS	Freilinger et al., 2009
Indian	0.200	0.632	0.335	0.033	212	IS	Banerjee et al., 2008
	0.213	0.616	0.341	0.043	232	CAD	Banerjee et al., 2009
	0.206	0.624	0.338	0.038	444		Mean for Indian
Saudi	0.231	0.607	0.324↓	0.069↑	929#	CAD	Alsmadi et al., 2009
Japanese	0.305	0.479	0.430	0.091	451	MI*	Ozaki et al., 2006
	0.302	0.485	0.427	0.089	2851	MI^	Ozaki et al., 2006
	0.315	0.469	0.433	0.098	2186	MI	Takashima et al., 2007
	0.320	0.463	0.427↓	0.110↑	1838	CAD	Hinohara et al., 2009
	0.345	0.445	0.422	0.133	1173#	CAD	Ikeda et al., 2012
	0.317	0.468	0.428	0.104	8499		Mean for Japanese
Chinese	0.320	0.455	0.450	0.095	2643	MI	Liu et al., 2009
Taiwanese	0.325	0.421	0.507↑	0.070↓	1097#	Population	Publication I
Korean	0.350	0.423	0.462	0.115	716	CAD	Hinohara et al., 2009
HCB	0.378	0.378	0.489	0.133	45	HapMap	NCBI resources
JPT	0.378	0.341	0.500	0.133	45	HapMap	NCBI resources

HapMap population data are given in italic, average values are underlined. Symbols (*) and (^) indicate respectively the data of preliminary and replicative studies performed by Ozaki with the colleagues and given in the same publication (Ozaki et al., 2006). Symbol ($^{\#)}$ indicates populations having genotype distribution deviated from HWE ($P < 0.05$) and arrows ↓ and ↑ indicate genotypes under- and over-represented respectively.

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., 2007; Trapina et al., 2009). Similar to the *PSMA3*, 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)).

2.3.2 Association of UPS genes polymorphisms with polygenic diseases

Disease-specific and common HLA and non-HLA genetic markers in susceptibility to rheumatoid arthritis, type 1 diabetes mellitus and multiple sclerosis

The common spectrum of autoimmune diseases affect the majority of tissues within the body, including pancreatic beta cells in type 1 diabetes (T1DM), myelin surrounding nerve axons in Multiple sclerosis (MS) and synovial joint antigens in Rheumatoid Arthritis (RA). The diseases are likely caused by a complex interaction between multiple HLA- and non-HLA related genes and environmental factors. The well documented co-clustering of autoimmune diseases within families and individuals, together with apparent sharing of number risk genes between the diseases suggests at least some common mechanisms of autoimmune development. The three diseases with possibly different manifestations were chosen for this review, the whole field being very vast. The fact all the three diseases with local rather than general symptoms (joints, pancreatic beta cells, myelin) possess common susceptibility factors stresses the importance of these findings.

Rheumatoid arthritis (RA) is one of the most common autoimmune diseases; it is chronic, systemic disease, manifesting as progressive, symmetrical deformation of joints, pain and inflammation. Around 1000 new cases of rheumatoid arthritis are diagnosed in Latvia each year. Both genetic and environmental factors play role in the susceptibility to this disease. Many genetic factors are known to increase risk for developing RA. The main genetic factor, involved in susceptibility and pathogenesis of RA is *HLA-DRB1* gene in the Major Histocompatibility complex (MHC) on chromosome 6. The main players in the pathogenesis of RA are cytokines; therefore polymorphisms in the regions that might affect expression of

the cytokine genes are extensively studied in this disease. Also anti-cytokine agents are successfully used as drugs for this disease. Genome-wide association studies have shown around 200 loci, which might be interesting for rheumatoid arthritis. These loci include genes, mainly involved in immune pesponse and antigen presentation, importance of which for RA is already known and extensively studied including the PTPN22 (protein tyrosine phosphatase expressed by the majority of cells belonging to the innate and adaptive immune systems), STAT4 (signal transducer and activators of transcription, nuclear import of STAT4 is linked to its tyrosine phosphorylation), CTLA4 (cytotoxic T lymphocyte antigen-4, a major negative regulatory molecule for T-cell activation), TRAF1/C5 (tumor necrosis factor-receptor associated factor 1/complement component 5), PADI4 (peptidyl-arginine-deiminase type IV), MIF (cytokine which counteracts the immunosuppressive properties of glucocorticoids) and many others (Ballina et al., 2008; Robinson et al., 2010; Raychaudhuri et al., 2010; Pasha et al., 2012; Chang et al., 2012), as well as others genes with yet unknown function. RA patients can be classified in different clinically significant groups according their genomic traits (ACPA (anti-citrullinated-protein-autoantibody) positive versus negative, (Seielstad et al., 2011). RA is very heterogeneous disease not only concerning the symptoms, genetic predisposition and antibody status, but also concerning variations in response to different therapeutic approaches. Several genes were found to affect the incidence of side effects and patients response to different drugs such as Methotrexate, these are *RFC1* (reduced folate carrier; (Giammarco et al., 2012), *ABCB1* (this gene encodes an efflux pump, ATP-binding cassette, sub-family B, member 1 (Takahashi et al., 2006), and *MTHFR* (methylenetetrahydrofolate dehydrogenase 1; Park et al., 2011; Lunt et al., 2011) genes. Some genes modify effect of anti TNF therapy: TNFA (tumour necrosis factor alpha; O'Rielly et al. 2009), interferon (IFN type 1) gene, promoter polymorphisms of interleukin IL-10 and other genes (Plant et al., 2011). Efficiency of Azatioprine therapy depends on TPMT gene, sulfasalazine - on NAT2 gene (Soejima et al., 2008). The RA susceptibility in Latvian population has been studied already (Mihailova et al. 2011), data on *TNFA* (rs1800629, rs361525), *IL6* (rs1800795), *IL18* (rs1946518, rs1946519), *IL10* (rs1800894, rs1800871 rs1800895, rs1800872), *IRF5* (rs3757385, rs2004640, rs10954213), *KLF12* (Krüppel-like factor 12; rs1324913) and *PTPN22* (rs2476601) associations with the disease are published.

Diabetes mellitus (DM) and its complications cause numerous health and social problems throughout the world. Number of DM patients constantly increases. Diabetes, characterized by persistent elevation of blood glucose levels (hyperglycaemia), occurs due to inadequate production of insulin (type 1 diabetes; T1DM), or resistance to endogenous insulin usually

associated with the metabolic syndrome and obesity (type 2 diabetes; T2D). Despite intensive glycaemic control, individuals with T1DM and T2DM are predisposed to developing vascular complications (reviewed in Sharma et al., 2012). DM and its complications are common n Latvia: in 2010 72654 DM patients were registered, 3891 with T1DM and 68217 with T2DM. Kidney malfunction like microalbuminuria was detected in 2737 patients, proteinuria – in 1092, renal failure – in 228, substitution therapy was indicated to 57 patients, kidney grafts were transplanted to 31 patient. Etiologically T1DM is an autoimmune disease both environmental and hereditary factors being important for susceptibility to it. Several studies have demonstrated a fundamental role for the HLA in the susceptibility of, or protection to T1DM. While HLA remains the strongest genetic risk factor, a number of novel gene variants associated with T1DM have been found through genome-wide studies, some of which have been linked to suspected environmental risk factors. However, jointly non-HLA risk alleles described up till now confer only a small additional risk compared to the effect of HLADR, HLADQ. Variants or polymorphisms in *UBASH3A* (a suppressor of T cell receptor signalling pathway) and *PTPN22* (the protein tyrosine phosphatase nonreceptor type 22) were shown as candidates for development of islet autoimmunity and T1DM when controlling for family history and in presence of the HLA-DR3/4-DQB1*0302 genotype. Polymorphisms in the insulin INS gene predicted development of T1DM. Although the effect of each individual gene is small, the combination of T1DM family history, the HLA-DR3/4-DQB1*0302 genotype, and the susceptibility variants of *PTPN22, UBASH2*, and INS were reported to increase 16-fold the risk of islet autoimmunity and 40-fold of T1DM.

Diabetic kidney disease, or diabetic nephropathy (DN), is a major complication of diabetes and the leading cause of end-stage renal disease (ESRD) that requires dialysis treatment or kidney transplantation. In addition to the decrease in the quality of life, DN accounts for a large proportion of the excess mortality associated with T1DM. The strongest association with DN as a primary phenotype was seen for an intronic SNP in the receptor tyrosine kinase *ERBB4* gene (Sandholm et al., 2012). DNA damage by endogenous free radicals and further ineffective repair of these lesions is considered to be one of the etiological factors of DM (Selvaraju et al., 2012; Tatsch et al., 2012). Oxidative stress and abnormal production of nitric oxide are of the main causes of increased DNA breakage in DM leading to development of its complications. Thus polymorphism of genes responsible for the free radical production and scavenging appear to be of special interest. Endothelial nitric oxide synthase (*eNOS*) polymorphisms have been reported to be strongly associated with DN risk, there are also reports on association of some alleles of inducible NO synthase (*iNOS)* gene with DM complications (Zhou et al., 2013; Uthra et al., 2011). Increased DNA breakage due

to oxidative stress determines importance of DNA repair genes (base excision repair gene *MUTYH,* X-ray repair cross complementing group 1XRCC1 and 8-Oxoguanine-DNA glycosylase hOGG1) in resistance or susceptibility to DM (Kasznicki et al., 2009; Chen et al., 2011).

Multiple sclerosis (MS) is most common, clinically extremely heterogeneous, chronic inflammatory disease of the CNS affecting about 2.5 million people around the world (2500 of them in the Latvia), presumably young adults, with onset usually at the second to fourth decade of life and, similarly to other autoimmune diseases, women being affected 3-4 times more frequent than men. (Compston et al., 2006). Pathologically, there are perivenular infiltrates of $Cd4^+$ and $CD8^+$ T cells in the CNS white matter and meninges with demyelinating lesions and loss of axons in both white and grey matter (Kornek et al., 2000). List of different MS subtypes includes the relapsing-remitting (RR), secondary progressive (SP) and primary progressive (PP), and within each subtype there is also considerable individual variation in disease course (Stys et al., 2012).

The cause of MS is not clear. Disease develops in genetically susceptible individuals with the contributions of environmental factors, such as infection, sunlight exposure, and vitamin D deficiency (Miller et al., 2011). There are also marked changes in systemic immune function with loss of regulatory T cell (Treg) function and increases in myelin-reactive $CD4^+$ T cells (Scheneider-Hohendorf et al., 2010; Venken et al., 2010; Dominguez-Villar et al., 2011).

Unbiased genome-wide association scans have identified susceptibility loci in regions containing genes with immune, co-stimulatory, signal transduction functions and vitamin D related, including, for example, the *CD6, CD25, CD40, CD58, HLA-A, HLA-B, HLA-DRB1, IL2RA, IL7R, IL12A, IRF8,* the Janus kinase (JAK)/signal transducer and activator of transcription *STAT3*, gene of the receptor for tumour necrosis factor-α *TNFRSF1A, CYP27B1* (1α-hydroxylase - an enzyme, which converts vitamin D to its active form) and other genes (Jager et al., 2009; IMSGC 2007, 2010; IMSGC and WTCCC2 2011; ANZgene 2009; Kofler et al., 2011). Osteopontin *OPN* gene may be involved in MS development and especially, progression and cytokines may be suggested as therapeutic target to counteract MS progression (Comi et al., 2012).

According to the ICSNPathway analysis (Identify candidate Causal SNPs and Pathway analysis) applied to the MS GWAS dataset (Zhang et al., 2011), 9 *HLA* and 7 *non-HLA* candidate SNPs and 5 and 10 candidate causal pathways respectively are mostly susceptible to MS including risk *HLA* loci: rs1802127 *(MSH5*), rs9277471 (*HLA-DPB1*), rs8084 and rs7192 (*HLA-DRA*), rs2072895 and rs2735059 (*HLA-F*), rs915669, rs915668 and rs1063320 (*HLA-*

G); and risk *non-HLA* loci: rs5896 (*F2*, prothrombin), rs8181979 (*SHC1*; (Src homology 2 domain containing) transforming protein 1), rs9297605 (*TAF2*, RNA polymerase II), rs669 (*A2 M;* M-type phospholipase A2 receptor), rs2228043 (*IL6ST*; interleukin 6 signal transducer), rs1061622 (*TNFRSF1B*), rs1801516 (*ATM*). The most strongly associated pathways were the rs1802127 to MSH5 to meiotic recombination and meiotic cell cycle and rs5896 to F2 to the transcriptional activation DNA-binding protein B from mRNA (Song et al., 2013).

Evidence underlines the importance of microRNAs (miRNAs) such as miR326, miR-323, miR-223, miR-23a, miR-15b and others in the MS pathogenesis (Fenoglio et al., 2012; Ridolfi et al., 2013). It was shown (Ridolfi et al., 2013) that rs1044165 (miR-223) likely acts as MS protective factor, while rs3745453 (miR-23a) seems to be a risk factor for MS. Extracellular miRNAs appear to be taken in consideration as a new source for both biomarker and risk factor identification and therapeutic drug discovery.

The RA, T1DM and MS common elements of genetic architecture.

Well documented co-clustering of autoimmune diseases within families and individuals are in good agreement with apparent sharing of a number of risk genes and polymorphic loci between particular diseases. At least the *IL7R-CAPSL* (interleukin 7 and calcyphosine-like protein) *CD226, Il2RA, HLA Class I, HLA Class II, CLEC16A, RGS1* (regulator of G protein signaling), *ZMIZ1* (zinc finger, MIZ-type containing 1), *TNFSF14* (tumor necrosis factor (ligand) superfamily, member 14) , *SOXB* (SRY (sex determining region Y)-box 3), *CLECL1* and *NFKB1* genes are susceptible to both the T1DM and MS. RA and MS share the susceptibility to variations in *VCAM1* (vascular cell adhesion molecule 1), *IL22RA2, PVT1, CLECL1* and *CD37* (IMSGC and WTCCC2, 2011; Brand and Gough, 2011; Zoledziewska et al., 2009). Despite large number of genes reported to be associated with each given disease, some polymorphic loci appear to be common for several diseases. These are, for example, the rs2104286 of the Il2RA sharing by MS and T1DM (Maier et al., 2009) and the rs10466829 of the CLECL1 sharing by RA, T1DM and MS (IMSGC and WTCCC2, 2011).

Functions of some non-HLA genes associated with autoimmune diseases are listed in Table 1. Several genes are involved in signaling pathways in T-lymphocytes (*PTPN22, STAT 4, CTLA-4, PADI4, MIF, PADI4*), apoptosis (*TRAF1/C5, TNFSF14*), calcium and G-protein-dependent signalling pathways (*CLEC16A, CAPSL, RGS1*) and other pathways.

Table 2.3.2.1. Non HLA genetic markers of autoimmune diseases

Abbreviated gene name	Full gene name	Function of the protein	Associated diseases	Sources
PTPN22	Protein tyrosine phosphatase	An inhibitor of T-cell activation, contributes to signaling cascades (TLR, TCR, BCR pathways) initiated in immune cells, including B cells and cells of the innate immune system.	RA, systemic lupus erythematosus (SLE) , T1DM, Hashimoto disease (HD), Graves'disease (GD) , vitiligo	Fousteri et al., 2013
STAT 4	Signal transducer and activators of transcription 4	Part of the JAK-STAT signalling pathway, expressed only spermatozoa, myeloid cells, and T lymphocytes. STAT 4 is activated by tyrosine phosphorylation in response to interleukin-12 (IL-12) treatment of T cells, involved in T helper cell function.	RA, SLE, Sjögren's syndrome, juvenile idiopathic arthritis (JIA), GD, myasthenia gravis	Reich, 2013; Korman et al., 2008
CTLA-4	Cytotoxic T-lymphocyte an tigen-4	Cell surface molecule involved in the regulation of signaling pathways affecting T-cell responses. Activation results in decreased T-lymphocyte activity and regulates the immune response.	RA, SLE, T1DM, Addison's disease (AD), Vitiligo, MS, HD	Fife, Bluestone, 2008; Kristiansen et al., 2000
TRAF1/C5	Tumor necrosis factor-receptor associated factor 1/complement component 5 locus	The TNF receptor associated factor 1 (TRAF1) is an adaptor protein, it is a TNF family member. TRAF1 is implicated in cell growth, proliferation, apoptosis, bone turnover, cytokine activation.	RA	Kurkó et al., 2013
PADI4	Peptidyl-arginine-deiminase type IV	The PADI4 enzyme is expressed in T and B cells, neutrophils, eosinophils, monocytes. It mediates the citrullination of histones (conversion of arginine residues to citrulline). Target for autoantibodies in RA.	RA	Kurkó et al., 2013; Anzilotti et al., 2010
MIF	Macrophage migration	A T cell derived cytokine, inhibits the random migration of	T1DM, RA	Sánchez-Zamora,

	inhibitory factor	macrophages in vitro and promotes macrophage accumulation during delayed-type hypersensitivity reactions.		Rodriguez-Sosa, 2014
UBASH3A	Ubiquitin associated and SH3 domain containing A	Belongs to T-cell ubiquitin ligand (TULA) family, facilitates growth factor withdrawal-induced apoptosis in T cells. An active phosphatase capable of dephosphorylating multiple tyrosine-phosphorylated proteins, suppressor of T cell receptor signalling pathway.	T1DM	Cerosaletti , Buckner, 2012; Tsygankov 2009.
STAT3	Signal transducer and activator of transcription 3 (acute-phase response factor)	STAT 3 protein is activated through phosphorylation by the receptor associated kinases in response to various cytokines and growth factors including IFNs, EGF, IL5, IL6, HGF, LIF and BMP2. Then it forms homo- or heterodimers that translocate to the cell nucleus where it act as transcription activator of a variety of genes in response to cell stimuli, and thus plays a key role in many cellular processes such as cell growth and apoptosis.	MS	De Jager et al., 2009
CLEC16A	C-type lectin domain family 16, member A	A transmembrane calcium-dependent (C-type) lectin-like receptor	MS, RA, T1DM	Zoledziewska et al., 2009
CAPSL	calcyphosine-like protein	calcium sensor and calcium signal modulator.	T1DM	IMSGC and WTCCC2, 2011
RGS1	Regulator of G protein signaling	Regulator of G protein signaling	T1DM, MS	IMSGC and WTCCC2, 2011
ZMIZ1	zinc finger, MIZ-type containing 1	A member of the PIAS (protein inhibitor of activated STAT) family of proteins. The encoded protein regulates the activity of various transcription factors, including the androgen receptor, Smad3/4, and p53. The encoded protein may also play a role in sumoylation.	T1DM, MS	IMSGC and WTCCC2, 2011
TNFSF14	tumor necrosis factor (ligand) superfamily,	A member of the tumor necrosis factor (TNF) ligand family. Functions as a	T1DM, MS	IMSGC and WTCCC2, 2011

	member 14	costimulatory factor for the activation of lymphoid cells, stimulates the proliferation of T cells, and triggers apoptosis of various tumor cells.		
VCAM1	vascular cell adhesion molecule 1	A cell surface sialoglycoprotein expressed by cytokine-activated endothelium, mediates leukocyte-endothelial cell adhesion and signal transduction.	T1DM, RA	IMSGC and WTCCC2, 2011
SOXB	SRY (sex determining region Y)-box 3	Member of SOX (SRY-related HMG-box) family of transcription factors involved in the regulation of embryonic development and in the determination of the cell fate.	T1DM, MS	IMSGC and WTCCC2, 2011
PVT1	Pvt1 oncogene (non-protein coding)	Oncogene, associated with several types of cancer and renal diseases.	T1DM, RA	(IMSGC and WTCCC2, 2011; Brand and Gough, 2011; Zoledziewska et al., 2009)

Coexistence of autoimmune diseases has been recently reviewed and statistically evaluated suggesting association of diseases and their common origin (Rojas-Villarraga et al., 2012). However the problem is only in the beginning of experimental evaluation and has to be very topical in genomic studies in nearest future.

2.4. Molecular genetics of juvenile idiopathic arthritic

Juvenile idiopathic arthritis (JIA) is the most common clinically heterogeneous chronic rheumatic disease in children (Ravelli et al., 2009). Onset-specific clinical features allow discrimination of seven JIA subtypes (Petty et al., 2004) with oligoarthritis (JIoA) and polyarthritis (JIpA) being the most frequent (Ravelli et al., 2009). The cause of JIA is complex, involving both environmental and genetic risk factors. The latter could include structural variations in both human leukocyte antigens and non-human leukocyte antigen candidate genes. (Prahalad et al., 2008; Thompson et al., 2010; Sivakumaran et al., 2011). Because of the pleiotropic effect, a frequent phenomenon in complex human traits and diseases (Sivakumaran et al., 2011) some loci of susceptibility may be shared with other autoimmune diseases (Thompson et al., 2010; Sivakumaran et al., 2011). An exceptional biological role for the ubiquitin proteasome system (UPS) in antigen processing and immune response, as suggested by Kloetzel (2001) has been increasingly supported this last decade

experimentally. A special form of proteasomes, thymoproteasomes, expressed exclusively in the cortex of the thymus and probably involved in positive selection of T cells, has recently been described. This indicates that the role of proteasomes in the immune response might be even more important (Sutoh et al., 2012; Egerer et al., 2002).

In patients with systemic autoimmune disease, the concentration of circulating proteasomes has been shown to be strongly increased (Egerer et al., 2002; Henry et al., 2011); the core 20S proteasome was identified as a target of the humeral autoreactive immune response.(Feist et al., 1996; 1999; Voigt et al., 2010; Arribas et al., 1991; Colmegna et al., 2008; Fissolo et al., 2008). The proteasomal inhibitor MG132 has been reported to reduce the severity of arthritis and reverse pain behaviour in arthritic rat models (Ahmed et al., 2010). Fine 14q13.2 microsatellite scanning revealed evidence of JIA association with variability in the region encompassing the *PSMA6* gene (Sjakste T, et al, 2010).

2.5. Molecular genetics of bronchial asthma

Asthma is a chronic inflammatory disease caused by complex gene-gene and gene-environment interactions with hyper-responsiveness to various nonspecific stimuli (Broide et al., 2001; Galli et al., 2008; Wu et al., 2012) being to a large extent genetically heterogeneous between human populations (Leung et al., 2014).

Ozaki et al. (2006) reported that the risk allele G of rs1048990 (*PSMA6*) might increase the expression of PSMA6 protein *in vivo* and *in vitro*, leading to enhanced inflammation through the activation of the NF-kB protein. Free NF-kB moves to the nucleus where it binds to target DNA elements and activate transcription of genes encoding proteins involved in immune responses, inflammation or cell proliferation. NF-kB could be considered as a coordinating element in the body's responses to situations of stress, infection or inflammation (Birrell et al., 2005; Edwards et al., 2009; Gagliardo et al., 2003). Thus, this polymorphism might be important for the development of chronic inflammation. However, interplay of multiple risk alleles and/or genotypes and primary driver of the disease remains still unclear.

Insufficient proteasome function was implicated in pathophysiology of various acute and chronic lung diseases and their complications (Albright et al., 2009; Majetschak et al., 2008; Sixt et al., 2010; Zemeckienė et al., 2013) and potentially could be a consequence of particular proteasomal genes structural variations. Early studies showed a possible linkage of chromosome 14 to asthma (Leung et al., 2014; Birrell et al., 2005).

Multiple studies including several GWAS analyses, indicated the 14q11-24 genome region as susceptible to asthma (Hakonarson et al., 2002; Malerba et al., 2001; Mansur et al., 2004; Munthe-Kaas et al., 2007; Tulah et al., 2012; Ungvári et al., 2012; Zhang et al., 2012).

This genomic region possess cluster of proteasomal genes including the *PSMA6*, *PSMC6* and *PSMA3* genes implicated previously in susceptibility to autoimmunity (Sjakste et al., 2004; 2010), type 2 diabetes mellitus (Sjakste et al., 2007), cardio-vascular disorders (Wang et al., 2013) and population adaptation to environment (Publication I). It appears that there is large potential for 14q proteasomal genes association studies to provide novel insights into bronchial asthma (BA) pathogenesis in particular human populations and in general.

2.6. Molecular genetics of obesity

The prevalence of obesity and overweight continues to raise worldwide causing serious health and social personal problems and substantial economic burden on societies. This complex multifactorial disease clearly implicated as a risk factor for many common diseases including diabetes, heart diseases, cancer and other health risk factors in all human populations (reviewed by Day & Loos 2011). Despite recent advances in the obesity genetics (reviewed by Day & Loos 2011; Loos 2009), further dissection of common inter-individual genetic backgrounds associated with and/or predisposed to obesity could bring more light on disease pathogenesis and assist to preventive strategies and more effective personalized disease treatment.

Structural variations in the proteasomal genes potentially could affect UPS efficiency through modulation of a particular gene expression, realization of gene and protein networks and metabolic processes that finally may influence predisposition to and/or development of obesity or overweight. It was shown that plasma ubiquitin and proteasome levels inversely correlated with body mass index (BMI) and could be considered as potential biomarkers for human obesity (Chang et al., 2009). Proteasome dysfunction mediates obesity-induced endoplasmic reticulum stress and insulin resistance in the liver (Otoda et al., 2013). Novel BMI susceptible locus had been recently identified on the chromosome 14 in Chinese population (Zhang et al., 2012). Decreased proteasomal activity causes age-related phenotypes and promotes the development of metabolic abnormalities (Tomaru et al., 2012). Significant association was found between obesity-associated phenotypes in Italians and genetic variants of the *PSMD9* gene encoding one of 26S proteasome non-ATP-ase regulatory subunits (Gragnoli et al., 2011). Mutation in the *PSMB8* gene encoding the immunoproteasome specific peptidase, has been reported to be associated with Japanese autoinflammatory syndrome with lipodystrophy (Kitamura et al., 2011). It was shown that plasma ubiquitin and proteasome levels inversely correlated with male body mass index in Southern Taiwan and Japanese population (Chang et al., 2009; Sakamoto et al., 2010) and proteasome dysfunction mediates obesity-induced endoplasmic reticulum stress and insulin

resistance in the liver (Otoda et al., 2013). Significant association was found between genetic variants of the *PSMD9* gene and obesity-associated phenotypes in Italians (Gragnoli et al., 2011). Mutation in the *PSMB8* gene has been reported to be associated with autoinflammatory syndrome with lipodystrophy in Japanese (Kitamura et al., 2011). Earlier was detected association of PSMA3 gene polymorphisms with susceptibility to obesity in Latvian children **(**Kupca et al., 2013)

2.7. Functional significance of SNPs

The role of heredity in susceptibility to diseases manifests itself on several levels of genome organisation: primary structure of DNA (genetical factor *per se*), interacting with various epigenetic factors like base modifications, modifications of chromatin packaging, etc. Therefore impact of genomic factors on human health should be analysed in complex, taking into account interactions of several levels of genome organisation.

The development of the different forms of cardiovascular, allergic, gastrointestinal and cancer diseases is determined by a complex interplay of environment factors with genetically determined features metabolism of a particular individual. Currently the traditional view that susceptibility to diseases is determined only by the interaction between the genes and the environment is being supplemented and expanded with new data about the key role of epigenetic reprogramming (Bird, 2002; Bruce et al., 2011; Hogg et al., 2012), heritable changes in gene expression that are not related to changes in the DNA sequence (Rodenhiser et al., 2006).

Initial studies in the field of epigenetics stressed the importance of genetic basis in development of multifactorial diseases. Environmental factors also influence epigenetical status of the organism, and interaction of both hereditary and environmental factors maintains "health" of the genome (Bird, 2002; Kellermayer et al., 2012). This explains existence of individual variations and uniqueness of cells, tissues and organs, in spite of the identity of the genetic information.

Many diseases are known to have a genetic component, but the epigenetic mechanisms underlying many conditions are still under study. Changes in the expression of genes within the body is characteristic of a significant number of disease, probably this is achieved via epigenetic mechanisms. These changes can be the cause of symptoms to the disease. Several diseases, especially cancer, have been suspected of selectively turning genes on or off in order to prevent the host's immune systems from destroying the tumorous tissues (Esteller et al., 2008).

The main epigenetic mediators assumed histone modifications, DNA methylation and non-coding RNAs (Vliet et al., 2007). It is known that the human genome contains 23,000 genes that must be expressed in cells of different specificities in exactly fixed time. This process depends largely on the chromatin structure. If chromatin is condensed, the genes are in inactivated state, if the chromatin is in open form and activated, it determines the activation of gene expression. This process is determined by reversible processes such as DNA methylation and histone modification. It is believed that the change in the activity of these processes leads to deregulation of the gene expression and can induce pathological processes (Miltenberger-Milteny et al., 2003; Sharp et al., 2004; Yang et al., 2012). The final category of epigenetic mechanism is regulatory RNA (MicroRNAs), the small, noncoding sequences that are involved in gene expression. Thousands of them are known, and the extent of their involvement in epigenetic regulation is an area of ongoing research (Prasanth, 2004).

Human single nucleotide polymorphisms (SNPs) represent the most frequent type of DNA variation in humans and many of them are believed to cause phenotypic differences between individuals. The investigation of SNP can help to understand the genetics of the human phenotype variations and especially the genetic basis of human complex diseases. Non-synonymous coding SNPs (nsSNPs) comprise a group of SNPs that, together with SNPs in regulatory regions, are believed to have the highest impact on phenotype (Ramensky et al., 2002).

The effect of SNPs involved in complex human phenotypes depends on many genetic and environmental components. In other words SNPs may comprise risk factors of having a specific phenotype in the statistical sense, and the effect of a particular SNP on phenotype might be seen only as a frequency difference between individuals that display the phenotype and unaffected controls.

The main point of disease gene identification should be functional analysis of the disease-associated allele for understanding of the molecular mechanism of causation of the disease phenotype. A possible way to overcome the problem of testing main numbers of SNPs, especially in the case of candidate gene studies, would be to determine SNPs according to their functional significance (Emahazion et al., 2001; Schork et al., 2000). This knowledge can be used to reduce the number of studied SNPs by focusing on specific genomic regions or gene sets, bioinformatics expertise may help to discriminate between neutral SNPs, which constitute the majority of genetic variation, and SNPs of likely functional importance.

Various approaches have been developed to identify variants that are likely to play an important biological role. Most of these approaches focus on the interpretation of coding or other SNPs in transcribed regions (Ng and Henikoff 2003; Adzhubei et al. 2010; Saccone et

al. 2010). The vast majority of associated SNPs identified in GWAS, however, are in non-transcribed regions, and it is likely that the underlying mechanism linking them to the phenotype is regulatory. A major challenge in the interpretation of GWAS results in a fact that SNPs located in close proximity in the genome tend to be in linkage disequilibrium (LD) with each other (The International HapMap Consortium 2005, 2007), and only a few SNPs per linkage disequilibrium region are measured on a given genotyping platform. Regions of strong linkage disequilibrium can be large, and SNPs associated with a phenotype have been found to be in perfect linkage disequilibrium with SNPs several hundred kilobases away.

SNPs that influence gene expression (expression quantitative trait loci, eQTLs) (Stranger et al. 2007; Schadt et al. 2008) have been shown to be significantly enriched for GWAS associations (Nicolae et al., 2010; Zhong et al. 2010). Several recent analyses of associated regions use these types of functional data in order to identify functional loci in individual diseases (Lou et al., 2009; Carvajal-Carmona et al., 2011; Harismendy et al., 2011; Paul et al., 2011). A recent study of chromatin marks in nine different cell lines produced a genome-wide map of regulatory elements and showed a twofold enrichment for predicted enhancers among the associated SNPs from GWAS (Ernst et al., 2011). These examples illustrate the power of statistical associations between a region of the genome and a phenotype.

Chromatin accessibility has been studied using DNase-seq, which led to the identification of 2.89 million DNase I–hypersensitive sites that may exhibit regulatory function. (The ENCODE Project Consortium 2004, 2007, 2011). DNase foot printing (Hesselberth et al., 2009; Boyle et al. 2011; Pique-Regi et al., 2011) was used to detect binding between proteins and the genome at a nucleotide resolution. ChIP-seq experiments were conducted for a total of 119 transcription factors and other DNA-binding proteins. Together these data provide a rich source of information that can be used to associate GWAS results with functional data.

High-throughput functional assays (chromatin immunoprecipitation assays) followed by sequencing (ChIP-seq) (Johnson et al., 2007; Robertson et al., 2007) and DNase I–hypersensitive site (Gross et al., 1988) identification by sequencing (DNase-seq) (Crawford et al., 2006; Boyle et al., 2008) can experimentally detect functional regions such as transcription factor binding sites (TFBS). Approaches based on known transcription factor binding motifs (Xu et al., 2009; Macintyre et al., 2010) have been successfully used to identify specific loci that have a functional role (Jarinova et al., 2009; Landers et al., 2009). The presence of SNPs in these regions leads to differences in transcription factor binding between individuals, as shows experimental evidences (Kasowski et al., 2010). A SNP that overlaps an experimentally detected TFBS and is in strong linkage disequilibrium with a SNP

associated with a phenotype can assume more likely play a biological role than other SNPs in the associated region for which there is no evidence of overlap with any functional data.

2.8. Perspectives for medical application.

The phenotypic biomarkers integrate of both genetic and non-genetic factors and can be used for clinical prognosis, and genotyping for specific SNPs will be useful in clinical diagnosis and prognostic assessment of patients. SNP markers are already being used in the diagnosis of a few diseases such as Wilson disease (Schmidt et al., 2007).

UPS components possess potential to be a therapeutic target for treatment of several diseases (Bedford et al., 2011). For example, the proteasome gene *PSMA6* polymorphism rs1048990 is associated with three different diseases: myocardial infarction (Wang et al., 2013; Heckman et al., 2013; Liu et al., 2009), type 2 diabetes (Liu et al., 2012; Barbieri et al., 2008), and coronary artery disease (Wang et al., 2013). It has recently been shown that mutations and polymorphisms in the immunoproteasome catalytic subunit *PSMB8* are associated with several inflammatory and autoinflammatory diseases including CANDLE syndrome (Torrelo et al., 2010), intestinal *M. tuberculosis* infection (Lv et al., 2011), and Nakajo-Nishimura syndrome (Arima et al., 2011). Polymorphisms in the insulin *INS* gene predicted development of T1DM (Wang et al., 2006.)

The putative role of *PSMA6* (8C>G) as survival factor in multiple myeloma has been shown for the first time: the G-allele was associated with worse 5-year survival, and it remained as an independent prognostic factor in multivariate analysis with a twofold higher risk to die compared to patients carrying the CC genotype. (Bachmann et al, 2010).

The proteasome may contribute to disease by increasing disease related liability in cells, and by resulting in reduced numbers of diseased cells. A distinct class of cancer-specific liabilities resulting from genome instability was recently reported (Nijhawan et al., 2012). Cells containing partial *PSMC2* copy number loss lack a proteasome complex composed of the protein product of *PSMC2, Rpt1,* and three other 19S subunits and eventually die after *PSMC2* suppression (Nijhawan et al., 2012).

In patients with rheumatoid (Egerer et al., 2002) and psoriatic (Henry et al., 2011) arthritis was shown to be substantially elevated the concentration of circulating proteasomes. Earlier, the 20S proteasome has been identified as a target of the humoral auto reactive immune response in patients with systemic inflammatory diseases including autoimmune myositis (Feist et al., 1996), primary Sjögren's syndrome (Feist et al., 1999), dilated cardiomyopathy (Voigt et al., 2010), systemic lupus erythematosus (Feist et al., 1996; Voigt et al., 2010; Colmegna et al., 2008), multiple sclerosis (Fissolo et al., 2008) and psoriatic

arthritis (Colmegna et al., 2008). The proteasomal inhibitor MG132 has been shown to reduce the severity of arthritis and reverse the pain behavior in the arthritic rat models (Ahmad et al., 2010). The immunoproteasome *PSMB9* codon 60HH variant was observed to have a reduced risk of developing multiple sclerosis in HLAA ∗02+ Italian females (Mishto et al., 2010). Results of treatment of autoimmune diseases using proteasome inhibitors have been successful in animal models (Paz et al., 2013; Fierabracci et al., 2013).

An array of natural and synthetic inhibitors of the proteolytic sites on the 20S proteasome have been developed both as research tools and as therapeutic agents (Nalepa et al., 2006; Goldberg et al., 2002; Gaczynska et al., 2005; Kisselev et al., 2001; Tsukamoto et al., 2006). Currently there are several inhibitors either approved or in clinical trials for the treatment of multiple cancers and strokes (Kisselev et al., 2001; 2006). These inhibitors have the capability of blocking protein degradation by the ubiquitine-proteasome pathway. They have also facilitated the discovery of numerous novel regulatory functions of UPS. The novel class of anticancer drug was represented by proteasome inhibitors (Richardson et al., 2006). Preclinical studies have showed that a dipeptidyl boronic acid, that is a selective inhibitor of the 26S proteasome, induces apoptosis and decreases proliferation, enhances the activity of radiation and chemotherapy and reverses hemoresistance in a variety of hematologic and solid malignancy models both *in vitro* and *in vivo* (Nawrocki et al., 2005). Dipeptidyl boronic acid was the first proteasome inhibitor to enter clinical trials and recently, received Food and Drug Administration (FDA) approval for the use of multiple myeloma and is being evaluated for the treatment of solid tumors (Nawrocki et al., 2005; Papandreou et al., 2004; Richardson et al., 2004). In preclinical animal models, selective proteasome inhibitors effectively suppress inflammatory arthritis and other inflammatory conditions (Palombella et al., 1998; Meng et al., 1999; Grisham et al., 1999; Bo Kim et al., 1999). In recent years, numerous novel drugs for the treatment of rheumatologic diseases have been successfully developed. The main approach has been to target cytokines, immune cells and their activation pathways (Olsen et al., 2004).

The ability of proteasomes to control the NF-*κ*B activity was shown in studies on proteasomal inhibitors. Silencing of the NF-*κ*B pathway through proteasomal inhibition has shown the promising results in the studies on various cell lines and small animal models with asthma (Elliott et al., 2003; Moutzouris et al., 2010). The proteasomal inhibitors were shown as potential drugs for anticancer therapy (Kalis et al., 2003). The clinical studies carried out with these inhibitors in subjects with cancer have also shown good results (Shah et al., 2002; Millward et al., 2012). Therefore, the research on proteasomes and their inhibitors could provide new therapeutic agents for the treatment of

various diseases, such as chronic inflammatory diseases, multiple sclerosis cancer, and other (Elliott et al., 2001).

3. REZULTS

PUBLICATIONS I; II; III; IV; V

PUBLICATION I

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

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.

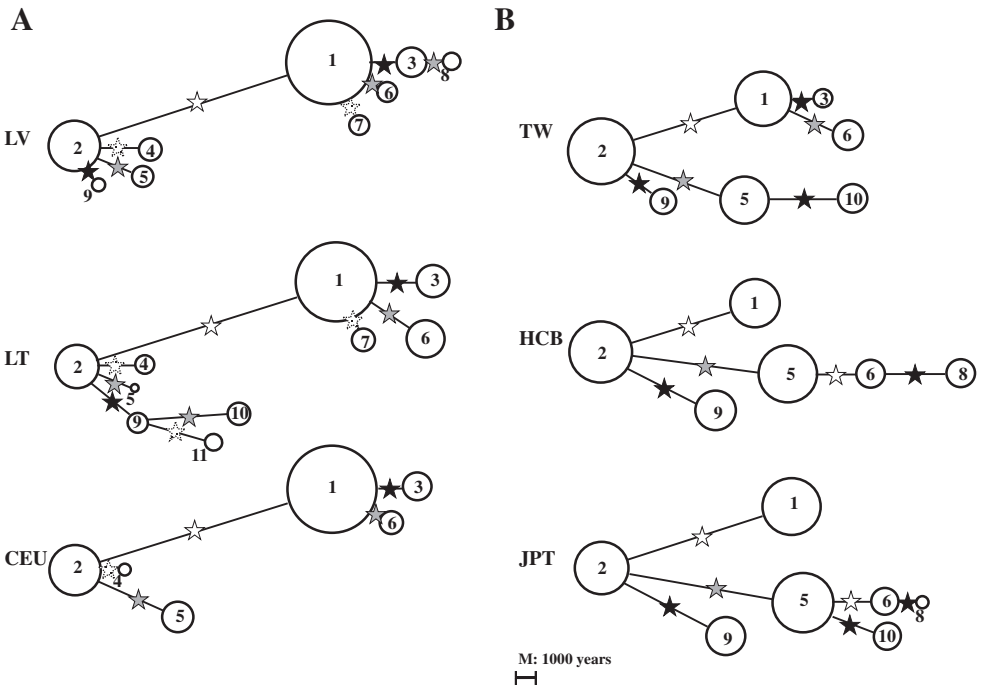


Fig. 1. Median joining network of *PSMA6*/*PSMC6*/*PSMA3* haplotypes. The network was generated using the Reduced Median algorithm of the Phylogenetic network software Fluxus 4.611 (http://www.fluxus-engineering.com). Panel A: LV — Latvians, LT — Lithuanians, CEU — NorthWestern Europeans; Panel B: TW — 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: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

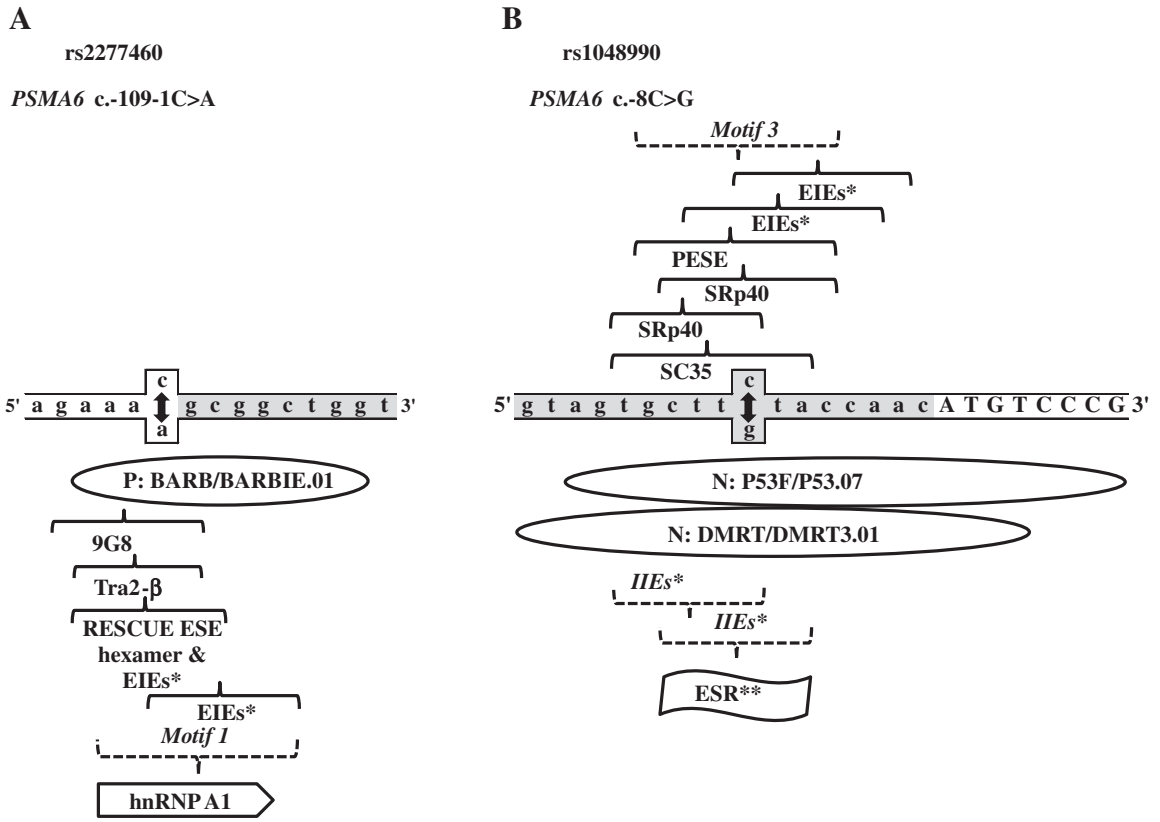


Fig. 2. Consequences of the rs2277460 (Panel A) and rs1048990 (Panel B) nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMA6* gene. Promoter and exon are coloured in white, 5′-UTR is coloured in grey; sequences of coding and noncoding genes' regions are presented 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.

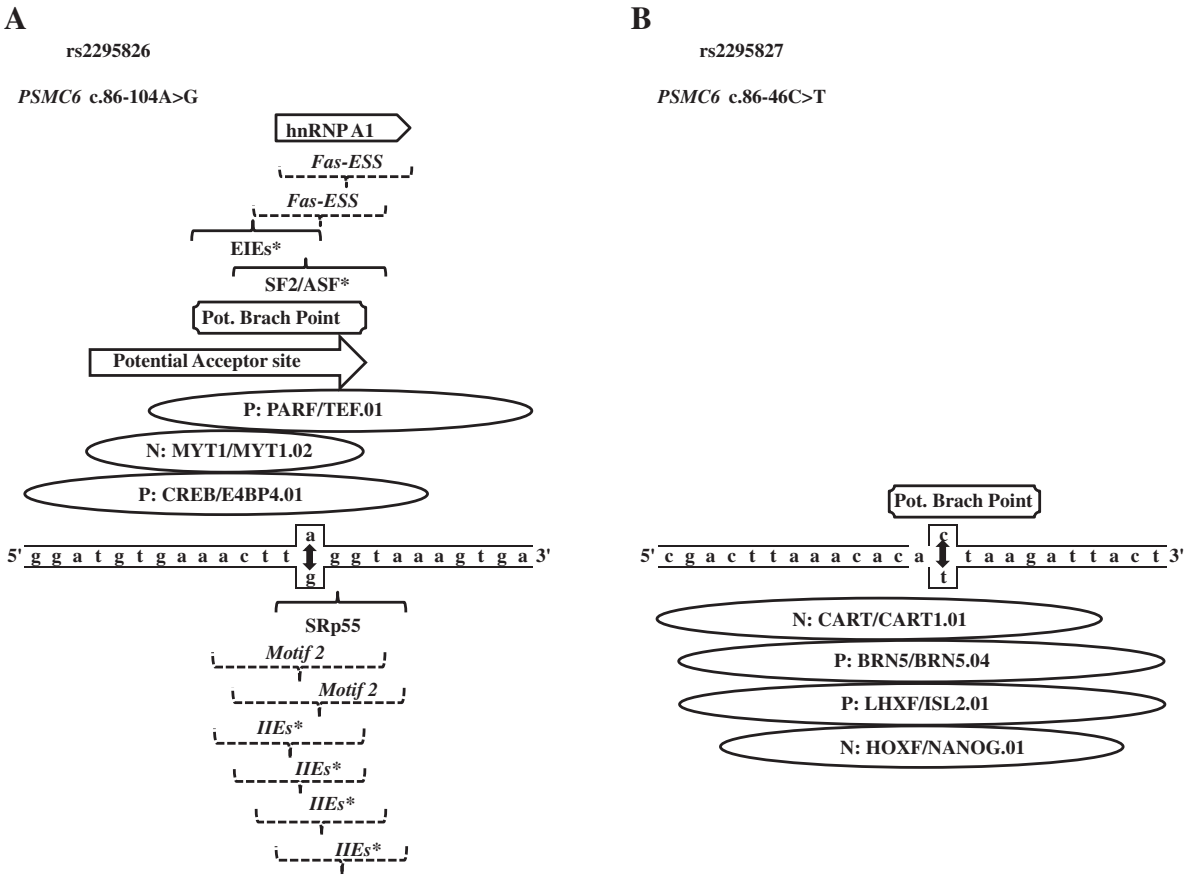


Fig. 3. Consequences of the rs2295826 (Panel A) and rs2295827 (Panel B) nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMC6* gene. Sequence features and other marks are given as described in Fig. 2. CREB/E4BP4.01–E4BP4, bZIP domain, transcription repressor; 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.

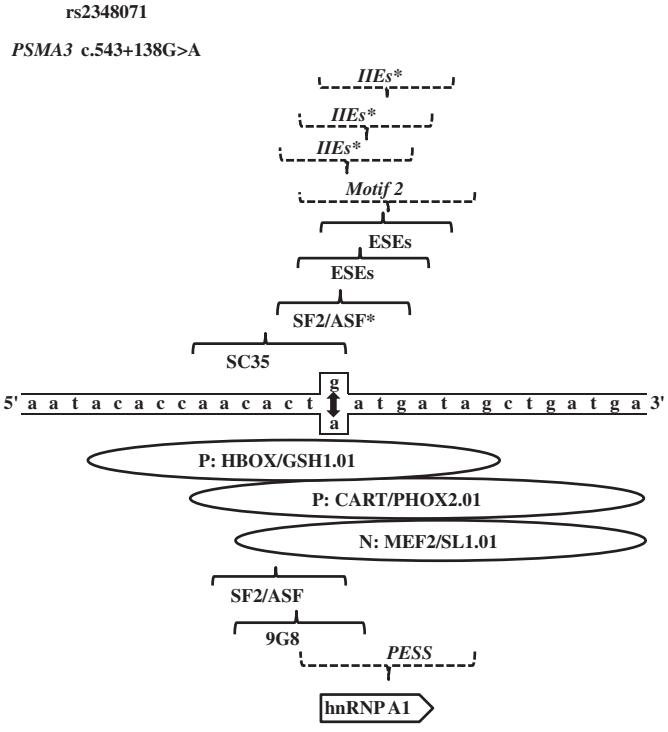


Fig. 4. Consequences of the rs2348071 nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMA3* gene. Sequence features and other marks are given as described in Fig. 2. HBOX/GSH1.01–Homeobox TF Gsh-1; Cart/PHOX2.01–Phox2a (ARIX) and Phox2b of cartilage homeoproteins 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.

PUBLICATION II

3.2. Juvenile Idiopathic Arthritis Subtype- and Sex-specific Associations with Genetic Variants in the *PSMA6/PSMC6/PSMA3* Gene Cluster

Pediatrics and Neonatology (2014) 55, 393–403

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://www.pediatr-neonatol.com

ORIGINAL ARTICLE

Juvenile Idiopathic Arthritis Subtype- and Sex-specific Associations with Genetic Variants in the *PSMA6/PSMC6/PSMA3* Gene Cluster

Tatjana Sjakste [a,*], Natalia Paramonova [a], Ingrida Rumba-Rozenfelde [b], Ilva Trapina [c], Olga Sugoka [a], Nikolajs Sjakste [b,c]

[a] *Institute of Biology, University of Latvia, Salaspils, Latvia*
[b] *Faculty of Medicine, University of Latvia, Riga, Latvia*
[c] *Latvian Institute of Organic Synthesis, Riga, Latvia*

Received Sep 17, 2013; received in revised form Dec 25, 2013; accepted Jan 30, 2014
Available online 27 May 2014

Key Words
genotype–sex interaction;
juvenile idiopathic arthritis;
plasma proteasome;
polymorphism;
PSMA3;
PSMA6;
PSMC6

Background: The ubiquitin proteasome system plays an exceptional biological role in the antigen processing and immune response and it could potentially be involved in pathogenesis of many immunity-related diseases, including juvenile idiopathic arthritis (JIA).
Methods: The *PSMB5* (rs11543947), *PSMA6* (rs2277460, rs1048990), *PSMC6* (rs2295826, rs2295827), and *PSMA3* (rs2348071) proteasomal genes were genotyped on JIA subtype- and sex-specific association; plasma proteasome levels was measured in patients having risk and protective four-locus genotypes and eventual functional significance of allele substitutions was evaluated *in silico*.
Results: Loci rs11543947 and rs1048990 were identified as disease neutral and other loci as disease susceptible ($p < 0.05$). The rs2277460, rs2295826, and rs2295827 loci had the strongest association with oligoarthritis [odds ratio (OR) = 2.024, 95% confidence interval (CI) 1.101–3.722; OR = 2.371, 95% CI 1.390–4.044; OR = 2.183, 95% CI 1.272–2.737, respectively), but the rs2348071 locus was associated with polyarthritis in females (OR = 3.438, 95% CI 1.626–7.265). A strong ($p < 0.001$) association was detected between the rs2277460/rs2295826/rs2295827/rs2348071 four-locus genotypes and the healthy phenotype when all loci were homozygous on common alleles (OR 0.439, 95% CI 0.283–0.681) and with the disease phenotype when the rs2348071 and the rs2295826 and/or rs2295827 loci were represented by risk genotypes simultaneously (OR 4.674, 95% CI 2.096–10.425). Rarely observed in controls, the double rs2277460/rs2348071 heterozygotes were rather frequent in affected males and

* Corresponding author. Genomics and Bioinformatics, Institute of Biology of the University of Latvia, Miera Str. 3, LV2169 Salaspils, Latvia.
E-mail address: tanja@email.lubi.edu.lv (T. Sjakste).

http://dx.doi.org/10.1016/j.pedneo.2014.01.007

more strongly associated with polyarthritis ($p < 0.05$). Haplotypes carrying the rare rs2295826/rs2295827 and rs2277460 alleles showed a strong ($p < 0.001$) association with oligo- and polyarthritis, respectively. The plasma proteasome level was found to be significantly higher in females having four-locus risk genotypes compared with protective genotypes ($p < 0.001$). Sequence affinity to transcription factors and similarity to splicing signals, microRNAs and/or hairpin precursors potentially depend on allele substitutions in disease susceptible loci.
Conclusion: We demonstrate for the first time evidence of a sex-specific association of *PSMA6/PSMC6/PSMA3* genetic variants with subtypes of JIA and plasma proteasome concentrations. Theoretical models of the functional significance of allele substitutions are discussed.

1. Introduction

Juvenile idiopathic arthritis (JIA) is the most common clinically heterogeneous chronic rheumatic disease in children.[1] Onset-specific clinical features allow discrimination of seven JIA subtypes,[2] with oligoarthritis (JIoA) and polyarthritis (JIpA) being the most frequent.[1] The cause of JIA is complex, involving both environmental and genetic risk factors. The latter could include structural variations in both human leukocyte antigens and non-human leukocyte antigen candidate genes.[3,4,5] Because of the pleiotropic effect, a frequent phenomenon in complex human traits and diseases,[5] some loci of susceptibility may be shared with other autoimmune diseases.[4,5]

An exceptional biological role for the ubiquitin proteasome system (UPS) in antigen processing and immune response, as suggested by Kloetzel,[6] has been increasingly supported this last decade experimentally. A special form of proteasomes, thymoproteasomes, expressed exclusively in the cortex of the thymus and probably involved in positive selection of T cells, has recently been described. This indicates that the role of proteasomes in the immune response might be even more important.[7,8] In patients with systemic autoimmune disease, the concentration of circulating proteasomes has been shown to be strongly increased[8,9]; the core 20S proteasome was identified as a target of the humeral autoreactive immune response.[10–15] The proteasomal inhibitor MG132 has been reported to reduce the severity of arthritis and reverse pain behavior in arthritic rat models.[16]

Highly conserved from an evolutionary standpoint, proteasomal genes appear to be subject to multiple trait purifying selection. Structural variations in proteasomal genes could potentially affect UPS efficiency through modulation of expression of a particular gene, realization of gene and protein networks and metabolic processes that may in the end influence predisposition to and/or development of autoimmune disorders.

The distribution of proteasomal genes over the human genome displays a tendency of clustering in chromosomes. Nine of the proteasomal genes have been localized in the long arm of chromosome 14, including two β (*PSMB5* and *PSMB11*) and two α (*PSMA3* and *PSMA6*) subunits of the core 20S proteasome, ATPases (*PSMC1* and *PSMC6*), 11S non-ATPase activators (*PSME1* and *PSME2*) and the *PSMA3P* pseudogene. The 14q11, 14q13, and 14q22-32 regions carrying the mentioned genes have been reported previously as potentially susceptible to autoimmune,[17–24] and other complex diseases in European and/or Asian populations.[Suppl 1–15] Fine 14q13.2 microsatellite scanning revealed evidence of JIA association with variability in the region encompassing the *PSMA6* gene.[24]

The aim of the current study was to genotype six single nucleotide polymorphisms (SNPs) belonging to the *PSMA6* (rs2277460 and rs1048990), *PSMA3* (rs2348071), *PSMB5* (rs11543947) and *PSMC6* (rs2295826 and rs2295827) proteasomal genes for subtype- and sex-specific association with JIA; to evaluate plasma proteasome levels in JIA patients of different multi locus genotypes, and to perform an *in silico* prediction of eventual functional consequences of nucleotide substitutions, including sequence affinity to transcription factors (TFs) and similarity to splicing signals and microRNAs.

2. Materials and methods

2.1. Case–control study

Patients were 174 JIA children (108 girls) receiving consultation at the outpatient clinic of P. Stradins Clinical University Hospital and Children Clinical University Hospital Clinic *Gailezers* in Riga, Latvia. JIA was diagnosed and assignment of the JIA patients to subgroups was carried out according to the criteria of the International League of Association for Rheumatology.[2] For association analysis both persistent and extended JIoA, and rheumatoid factor-negative and rheumatoid factor-positive JIpA subgroups were combined in JIoA and JIpA groups of 107 and 55 patients, respectively. Twelve other patients were diagnosed as having systemic ($n = 9$), enthesitis-related ($n = 2$), and psoriatic ($n = 1$) arthritis.

The control group was represented by 191 (117 women) patents of Riga Bikernieki Hospital admitted with a diagnosis of trauma and not diagnosed as having any autoimmune and/or cardiovascular disorders, type 2 diabetes mellitus (T2DM), or obesity.

Informed consent was obtained from all the study participants or their parents. The study was approved by the Central Medical Ethics Committee of the Republic of Latvia Ministry of Health.

2.2. Marker choice

Due to limited data on the genetic diversity and susceptibility to diseases of proteasomal genes, several criteria were taken into account in choosing markers. These included the existence of previously reported findings on locus association with human health status, locus allele-specific potential to be functionally significant, locus variability in Latvians, Hardy–Weinberg expectations and others concerning mainly a genotyping technology. The rs2277460 and rs1048990 of the *PSMA6*, rs2295826 and rs2295827 of the *PSMC6* and rs23480071 of the *PSMA3* were previously studied on disease susceptibility[Suppl 1–12,14–16] and/or genetic diversity[Suppl 17]; rs11543947 of the *PSMB5* gene was previously genotyped on genetic diversity only in HapMap populations. All loci fit all other mentioned criteria of marker choice.

2.3. DNA extraction and genotyping

DNA was extracted using a kit for genomic DNA extraction from nucleated blood cells (Fermentas, Vilnius, Lithuania). Genotyping methods and primer sequences are indicated in Supplementary Table 1. Basic PCR was performed with DreamTaq polymerase (Fermentas) using the following parameters: 94°C for 5 minutes; then 35–40 cycles of 94°C for 45 seconds, appropriate annealing temperature (55–61°C) for 45 seconds, 72°C for 45 seconds and a final extension step at 72°C for 7 minutes. DNA digestion by restriction enzymes was performed according to the manufacturer's protocols (Fermentas).

Amplified and digested products were analyzed by electrophoresis in 1–3% agarose gel for all markers. For quality control, 16 randomly chosen samples for each marker were genotyped in duplicate in different experiments. 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 3130*xl* Genetic Analyzer. Alleles and genotype frequencies for the rs11543947 (ss69150930), rs2277460 (ss24557113), rs1048990 (ss35076445), rs2295826 (ss3239727 and ss69157456), rs2295827 (ss23619651) and rs2348071 (ss3302481) were obtained for HapMap-CEU (NorthWestern European), YRI (Yoruba), JPT (Japanese), and HCB (Han Chinese) populations from publicly available dbSNP (build 13) entries at NCBI (http://www.ncbi.nlm.nih.gov/snp/). Loci description and nucleotide numbering are given according to the recommended nomenclature system (http://www.genomic.unimelb.edu.au/mdi/mutnomen/recs.html). Sequence information for the chromosome 14 GRCh37.p5 assembly (NCBI reference sequence: NC_000014.8) was used for loci description, nucleotide numbering, and primer design using the Primer 3.0 program.

2.4. Measurement of plasma proteasome concentration

Plasma samples were available only for 23 JIA patients. These plasma samples were obtained from patients randomly chosen for plasma sampling during development of the DNA collection in the JIA study. Therefore, preliminary information on genotypes of these patients did not exist at the moment of the sampling. Blood was harvested on citrate anticoagulant, and plasma stored at −80°C. Plasma proteasome concentration was measured in triplicate for each sample using a standard 20S/26S Proteasome ELISA kit (BML-PW0575; ENZO Life Sciences, Farmingdale, NY, USA) according to the manufacturer's protocols. Absorbance was read at 450 nm using a UV-Vis spectrometric plate reader. Results were expressed as concentration of proteasome protein in ng/mL determined by interpolation for the absorbance value using the generated 20S proteasome standard curve.

2.5. Data analysis

Documenting personalized genotyping data allowed determination of rs11543947/rs2277460/rs1048990/rs2295826/rs2295827/rs2348071 six-locus genotype (6-LG) of each individual participant of the study. The 6-LGs, rs2277460/rs2295826/rs2295827/rs2348071 four-locus genotypes (4-LGs), observed haplotypes, single locus genotypes (SLGs), and allele frequencies were estimated by direct counting of genetic variants. Inferred haplotypes prediction, haplotype sorting, estimation of the linkage disequilibrium and probability of recombination were performed using the DnaSP software version 5.10.1 online tool at http://www.ub.es/dnasp.[Suppl 18] Both the two-tailed Fisher's exact test and the χ^2 test were applied to evaluate the linkage between the rs2295826 and rs2295827 polymorphic sites at three *p*-value levels ($p < 0.05$; $p < 0.01$; $p < 0.001$). The Bonferroni correction included in the DnaSP analysis was taken into account to support the significance of the revealed disequilibrium ($\alpha' = 0.05$).

Deviation from the Hardy–Weinberg equilibrium and differences between cases and controls in allele, genotype and haplotype frequencies were evaluated by χ^2 and Cochran–Armitage trend test using XLSTAT 2013 software for Windows. Genetic models for every individual locus were designed according to Lewis.[Suppl 19] Contingency tables were 2×3 for the AA, AB, BB genotypes in the general model; 2×2 for the AA, AB+BB and AA+AB, BB, and AB, AA+BB genotypes in the dominant, recessive, and over dominant models, respectively and A and B alleles in the multiplicative model where A is the major allele and B is the minor allele. Using an additive model, the AA, AB, and BB genotype distribution was analysed using the Cochran–Armitage test for trend. An odds ratio (OR) > 2 and < 0.5 was considered to be clinically significant. Stratification was performed by JIoA and JIpA ILAR subtypes and by sex.

Levels of plasma proteasome were expressed as mean ± standard error of the mean for each sample to show the variability associated with the estimation, and as mean ± standard deviation to characterize the spread of a data set within the groups. Both standard error of the mean and standard deviation were calculated using the online NCalculators (http://ncalculators.com/). Differences in plasma proteasome levels between the groups were estimated by nonparametric Mann–Whitney and/or Kruskal–Wallis tests using XLSTAT 2013 software. Results were considered to be of nominal statistical significance at $p < 0.05$, moderate statistical significance at $p < 0.002$, and strong statistical significance at $p < 0.001$.[Suppl 20]

2.6. SNP functional analysis in silico

An eventual functional significance of the SNPs showing evidence of association was analyzed *in silico* on sequence similarity to transcription factors binding sites (TFBSs) using Genomatix software, MatInspector, Release 7.4 online tool, at www.genomatix.de.[Suppl 21] Only parameters with core/matrix similarity of > 1.000/0.800 were taken into account. Splicing signals were predicted by Human Splicing Finder Version 2.4 (http://www.umd.be/HSF)[Suppl 22] with standard threshold values for branch point, donor and acceptor splice sites, enhancer, silencer, heterogeneous nuclear ribonucleoprotein (hnRNP) and other splicing motifs. Sequence similarity to mature microRNAs and hairpin precursors was evaluated, and microRNA targets prediction was done using miRBase (http://www.mirbase.org/index.shtml)[Suppl 23] and miRNAMap (http://mirnamap.mbc.nctu.edu.tw/index.php)[Suppl 24] online tools, respectively.

3. Results

3.1. Genotyping results and single locus association

In both case and control cohorts, the genotyping call rate was 100% and all six markers were found to be in Hardy–Weinberg equilibrium. Allele and genotype spectrum and distributions in Latvians were found to be similar to those of other Europeans (CEU) for all SNPs studied and to the Yoruba population (YRI) for the rs2348071; however, Latvians differ from YRI for resting loci and for all loci from Japanese (JPT) and Han Chinese (HCB) populations (Supplementary Table 2).

The rs11543947 and rs1048990 markers showed similar levels of variation in controls and JIA patients without significant differences between the JIoA and JIpA subtypes and females and males. These markers were considered to be JIA neutral, while other markers were found to be JIA susceptible (Table 1).

The rs2277460 showed nominal association ($P < 0.05$) with JIA, with highest risk effect for JIoA [OR = 2.024, 95% confidence interval (CI) 1.101–3.722]. Alleles and genotypes frequencies of the rs2295826 and rs2295827 were the same between each other in the controls and slightly different in JIoA females with an $r^2 = 0.936$ and D′ of 1.000 [this result is similar to data obtained for Tuscans in Italy (HapMap TSI): $r^2 = 0.923$], suggesting the existence of rare GC haplotype in JIA patients. Both markers were found to be in moderate ($p < 0.002$) association with JIoA (OR = 2.371, 95% CI 1.390–4.044; and OR = 2.183, 95% CI 1.272–2.737 for the rs2295826 and rs2295827 risk genotypes, respectively). Moderate association was also detected for the rs2348071 heterozygous genotype ($p < 0.002$) with risk effect for JIpA in the combined cohort

Table 1 Data on the significant associations between single locus variations and JIA.

Groups of comparison		Marker ID	Genetic model	Risk factor (allele or genotype)	Risk factor number (%)		*p*	OR	CI
Group 1	Group 2				Group 1	Group 2			
JIA (174)	C (191)	rs2277460	Multiplicative	a: A	40 (11.49)	25 (6.54)	<0.05	1.855	1.104–3.114
			Dominant	g: CA	40 (22.99)	25 (13.09)	<0.05	1.982	1.149–3.419
		rs2295826	Multiplicative	a: G	64 (18.39)	40 (10.47)	<0.05	1.927	1.262–2.942
			Dominant	g: AG + GG	55 (31.61)	36 (18.85)	<0.05	1.990	1.230–3.210
		rs2295827	Multiplicative	a: T	61 (17.53)	40 (10.47)	<0.05	1.817	1.186–2.784
			Dominant	g. CT + TT	53 (30.46)	36 (18.85)	<0.05	1.886	1.163–3.057
		rs2348071	Overdominant	g: AG	87 (50.00)	66 (34.55)	<0.05	1.894	1.245–2.882
JIA-F (108)	C-F (117)	rs2295826	Multiplicative	a: G	46 (21.30)	30 (12.82)	<0.05	1.840	1.116–3.033
			Dominant	g: AG + GG	40 (37.04)	27 (23.07)	<0.05	1.932	1.086–3.432
		rs2295827	Multiplicative	a: T	43 (19.91)	30 (12.82)	<0.05	1.690	1.020–2.801
			Dominant	g. CT + TT	38 (35.18)	27 (23.07)	<0.05	1.810	1.013–3.232
		rs2348071	Overdominant	g: AG	57 (52.78)	40 (34.19)	<0.05	**2.151**	1.261–3.672
JIoA (107)	C (191)	rs2277460	Multiplicative	a: A	25 (11.68)	25 (6.54)	<0.05	1.889	1.061–3.362
			Dominant	g: CA	25 (23.36)	25 (13.09)	<0.05	**2.024**	1.101–3.722
		rs2295826	Multiplicative	a: G	45 (21.03)	40 (10.47)	**<0.001**	**2.277**	1.435–3.612
			Dominant	g: AG + GG	38 (35.51)	36 (18.84)	**<0.002**	**2.371**	1.390–4.044
		rs2295827	Multiplicative	a: T	42 (19.63)	40 (10.47)	**<0.002**	**2.088**	1.308–3.333
			Dominant	g. CT + TT	36 (33.64)	36 (18.84)	**<0.002**	**2.183**	1.272–2.737
		rs2348071	Overdominant	g: AG	50 (46.73)	66 (34.56)	<0.05	1.661	1.027–2.687
JIoA-F (63)	C-F (117)	rs2295826	Multiplicative	a: G	31 (24.60)	30 (12.82)	<0.05	**2.219**	1.275–3.861
			Dominant	g: AG + GG	26 (41.27)	27 (23.07)	<0.05	**2.342**	1.216–4.511
		rs2295827	Multiplicative	a: T	28 (22.22)	30 (12.82)	<0.05	1.943	1.105–3.416
			Dominant	g. CT + TT	24 (38.09)	27 (23.07)	<0.05	1.943	1.105–3.416
JIpA (55)	C (191)	rs2348071	Overdominant	g: AG	32 (58.18)	66 (34.56)	**<0.002**	**2.635**	1.434–4.841
JIpA-F (39)	C-F (117)	rs2348071	Overdominant	g: AG	25 (64.10)	40 (34.19)	**<0.002**	**3.438**	1.626–7.265

$p < 0.002$ and odds ratio > 2 are indicated in bold.
JIA = juvenile idiopathic arthritis; JIoA = juvenile idiopathic oligoarthritis; JIpA = juvenile idiopathic polyarthritis; C = control; F = female; a = risk allele; g = risk genotype.

Table 2 Four-loci genotypes (4-LGs) presentation and results of significant associations with juvenile idiopathic arthritis (JIA).

4-LG configurations					4-LGs number (%) in the groups and association results											
No.	Genotype of individual locus				Controls			JIA			JIoA			JIpA		
	PSMA6	*PSMC6*		*PSMA3*	All $n = 191$	F $n = 117$	M $n = 74$	All $n = 174$	F $n = 108$	M $n = 66$	All $n = 107$	F $n = 63$	M $n = 44$	All $n = 55$	F $n = 39$	M $n = 16$
	L1	L2	L3	L4												
$1^{P;Ref}$	CC	AA	CC	GG or AA	86 (45.03)	52 (44.44)	34 (45.95)	46 (26.44)	24 (22.22)	22 (33.33)	32 (29.91)	17 (26.98)	15 (34.09)	11 (20.00)	6 (15.38)	5 (31.25)
					Association in the groups:	*p*	OR	95% CI								
					JIA vs. C	**<0.001**	**0.439**	0.283—0.681								
					JIA-F vs. C-F	**<0.001**	**0.357**	0.200—0.637								
					JIoA vs. C	<0.05	0.521	0.316—0.859								
					JIoA-F vs. C-F	<0.05	0.462	0.239—0.859								
					JIpA vs. C	**<0.001**	**0.305**	0.150—0.620								
					JIpA-F vs. C-F	**<0.002**	**0.227**	0.091—0.568								
2^{N}	CC	AA	CC	**GA**	48 (25.13)	28 (23.93)	20 (27.03)	44 (25.29)	30 (27.78)	14 (21.21)	20 (18.69)	12 (19.05)	8 (18.18)	19 (34.55)	15 (38.46)	4 (25.00)
3^{N}	CC	**AG** **GG**	**CT** **TT**	GG or AA	22 (11.52)	18 (15.39)	4 (5.41)	19 (10.92)	14 (12.96)	5 (7.58)	11 (10.28)	7 (11.11)	4 (9.09)	7 (12.73)	6 (15.38)	1 (6.25)
4^{N}	**CA**	AA	CC	GG or AA	16 (8.38)	7 (5.98)	9 (12.16)	16 (9.20)	9 (8.33)	7 (10.61)	9 (8.41)	6 (9.52)	3 (6.82)	4 (7.27)	2 (5.13)	2 (12.50)
5^{R}	CC	**AG** **GG** **GG**	**CT** **TT** CC	**GA**	10 (5.24)	7 (5.98)	3 (4.05)	25 (14.37)	19 (17.59)	6 (9.09)	19 (17.76)	14 (22.22)	5 (11.36)	6 (10.91)	5 (12.82)	1 (6.25)
					Association in the groups:	*p*	OR	CI								
					JIA vs. C	**<0.001**	**4.674**	2.096—10.425								
					JIA-F vs. C-F	**<0.001**	**5.881**	2.231—15.500								
					JIoA vs. C	**<0.001**	**5.106**	2.179—11.968								
					JIoA-F vs. C-F	**<0.001**	6.118	2.175—17.210								
					JIpA vs. C	<0.05	4.691	1.477—14.897								
					JIpA-F vs. C-F	<0.05	6.190	1.574—24.343								
6^{R}	**CA**	AA	CC	**AG**	5 (2.62)	3 (2.57)	2 (2.70)	13 (7.47)	5 (4.63)	8 (12.12)	8 (7.48)	2 (3.17)	6 (13.64)	5 (9.09)	3 (7.69)	2 (12.50)
					Association in the groups:	*p*	OR	CI								
					JIA vs. C	<0.05	4.861	1.625—13.940								
					JIA-M vs. C-M	<0.05	6.182	1.370—27.895								

(*continued on next page*)

Table 2 (*continued*)

No.	4-LG configurations: Genotype of individual locus				4-LGs number (%) in the groups and association results											
					Controls			JIA			JIoA			JIpA		
	PSMA6	*PSMC6*		*PSMA3*	All	F	M	All	F	M	All	F	M	All	F	M
	L1	L2	L3	L4	$n = 191$	$n = 117$	$n = 74$	$n = 174$	$n = 108$	$n = 66$	$n = 107$	$n = 63$	$n = 44$	$n = 55$	$n = 39$	$n = 16$
					JIoA vs. C	<0.05	4.300	1.367–13.524								
					JIoA-M vs. C-M	**<0.002**	6.800	1.405–32.913								
					JIpA vs. C	**<0.002**	6.800	1.405–32.913								
7	CA	AG	CT	GG or AA	1	—	1	6	4	2	5	4	1	1	—	1
		GG	TT		(0.51)		(1.35)	(3.44)	(3.70)	(3.03)	(4.67)	(6.36)	(2.27)	(1.81)		(6.25)
8	CA	AG	CT	GA	3	2	1	5	3	2	3	1	2	2	2	—
					(1.57)	(1.71)	(1.35)	(2.87)	(2.79)	(3.03)	(2.80)	(1.59)	(4.55)	(3.64)	(5.13)	

In statistical analysis the 4-LG1 frequency was compared to sum of all other 4-LGs frequencies; protective 4-LG1 was considered as reference genotype in the risk 4-LGs identification. [P] protective JIA genotype; [R] JIA risk genotype. Risk single locus genotypes are given in bold and underlined.
JIA = juvenile idiopathic arthritis; JIoA = juvenile idiopathic oligoarthritis; JIpA = juvenile idiopathic polyarthritis; F = female; M = male; L1 = rs2277460 locus; L2 = rs2295826 locus; L3 = rs2295827 locus; L4 = rs2348071 locus.

(OR = 2.635, 95% CI 1.434–4.841) and JIpA females (OR = 3.438, 95% CI 1.626–7.265).

In both the control and case groups, risk genotypes were more frequent in males for the rs2277460 locus and in females for the rs2295826 and rs2295827 loci; the rs2348071 heterozygous risk genotype was more frequent in females than in males in JIpA patients (Supplementary Table 2).

3.2. Identification of the risk/protective 4-LGs

Personalized genotyping data documentation allowed analysis of the spectrum and frequencies in the groups of the 4-LGs rs2277460/rs2295826/rs2295827/rs2348071 (Table 2) composed from loci individually susceptible to disease (Table 1). Nineteen 4-LGs observed in both case and control groups were classified by eight categories according to presence/absence of the risk SLG listed in Table 1.

The 4-LG1, having a no risk SLG, was the most frequent in controls (45%) with similar presentation in males and females, but it was significantly less frequent in JIA patients of both JIoA and JIpA subtypes (about 29% and 20% respectively) and appears to be JIA protective with a strong level of association ($p < 0.001$) with healthy phenotype in common cohort and females. The 4-LG5 unites three configurations, all having the rs2348071 risk genotype in combination with risk genotype at the rs2295826 and/or rs2295827 loci. The 4-LG5 was approximately three times more frequent in JIA than in controls and two times more frequent in females than in males of all JIA subtypes. This genotype showed a strong association ($p < 0.001$) with JIA in common and JIoA cohorts and the female phenotype. Rarely observed in controls and JIA females (< 5%), the rs2277460/rs2348071 double heterozygotes (4-LG6) were rather frequent in JIA males (> 12%) and showed an association ($p < 0.002$) with JIpA in common cohort and with JIoA in males.

The 4-LG2, 4-LG3, and 4-LG4 genotypes were observed with similar frequencies in cases and controls and were considered JIA neutral. The two remaining (4-LG7 and 4-LG8) genotypes were rare in controls and only slightly more frequent in JIA patients.

3.3. Four loci haplotype analysis

Table 3 provides information on the four-loci haplotype (4-LH) spectrum and frequencies in the groups. Using the assumption of random assortment of alleles, 24 haplotype configurations were expected for four two-allele loci; however, only 10 variants were identified in cases and controls taken together, and all of them are implicated from the 4-LGs homozygous at all four loci and/or genotypes being heterozygous only at one locus. The 4-LH1–4-LH6 haplotypes were observed in both controls and case groups; the 4-LH7–4-LH10 were identified only in JIoA females. The 4-LH1 (C-C-A-G) having the common alleles at all four loci was found to be the most frequent in all groups and used as reference haplotype in association analysis. The 4-LH4 (C-G-T-A) having the risk alleles at the rs2295826 and rs2295827 and the rs2348071 minor allele A was found to be in strong association ($p < 0.001$) with JIA including both JIoA and JIpA subtypes in female and male cohorts. The 4-LH5 (A-C-A-G)

Table 3 Four-loci haplotypes (4-LHs) presentation and data on significant associations with juvenile idiopathic arthritis (JIA).

Group		C *n* = 191	C-F *n* = 117	C-M *n* = 74	JIA *n* = 174	JIA-F *n* = 108	JIA-M *n* = 66	JIoA *n* = 107	JIoA-F *n* = 63	JIoA-M *n* = 44	JIpA *n* = 55	JIpA-F *n* = 39	JIpA-M *n* = 16
4-LG		**Number (%)**											
Full homozygote		89 (46.60)	54 (46.15)	35 (47.30)	49 (28.16)	25 (23.15)	24 (36.36)	34 (31.78)	17 (26.98)	17 (38.64)	12 (21.82)	7 (17.95)	5 (31.25)
Single locus heterozygote		65 (34.03)	36 (30.77)	29 (39.19)	66 (37.93)	44 (40.74)	22 (33.33)	34 (31.78)	23 (36.51)	11 (25.00)	24 (43.64)	17 (43.59)	7 (43.75)
Multiple loci heterozygote		37 (19.37)	27 (23.08)	10 (13.51)	59 (33.91)	39 (36.11)	20 (30.30)	39 (36.45)	23 (36.51)	16 (36.36)	19 (34.55)	15 (38.46)	4 (25.00)
Haplotype	**Loci**	**Number (%)**											
	1—2—3—4	382	234	148	348	216	132	214	126	88	110	78	32
4-LH-1^{Ref}	C-A-C-G	231 (60.47)	147 (62.82)	84 (56.76)	178 (51.15)	109 (50.46)	69 (52.27)	117 (54.67)	68 (53.97)	49 (55.68)	53 (48.18)	38 (48.72)	15 (46.88)
4-LH-2	C-A-C-A	86 (22.51)	45 (19.23)	41 (27.70)	72 (20.69)	45 (20.83)	27 (20.45)	32 (14.95)	19 (15.08)	13 (14.77)	27 (24.55)	19 (24.36)	8 (25.00)
4-LH-3	C-**G-T**-G	29 (7.59)	19 (8.12)	10 (6.76)	19 (5.46)	13 (6.02)	6 (4.55)	12 (5.61)	9 (7.14)	3 (3.41)	7 (6.36)	4 (5.13)	3 (9.38)
4-LH-4^{R}	C-**G-T**-A	11 (2.88)	11 (4.70)	—	37 (10.63)	25 (11.57)	12 (9.09)	25 (11.68)	14 (11.11)	11 (12.50)	11 (10.00)	10 (12.82)	1 (3.13)
		Association in the groups	*P*	OR	CI								
		JIA vs. C	**<0.001**	**4.365**	2.192—8.693								
		JIoA vs. C	**<0.001**	**4.487**	2.161—9.318								
		JIpA vs. C	<0.05	4.358	1.826—10.401								
4-LH-5	**A**-A-C-A	15 (3.93)	8 (3.42)	7 (4.73)	12 (3.45)	5 (2.31)	7 (5.30)	11 (5.14)	4 (3.17)	7 (7.95)	—	—	—
4-LH-6^{R}	**A**-A-C-G	10 (2.62)	4 (1.71)	6 (4.05)	22 (6.32)	11 (5.09)	11 (8.33)	9 (4.21)	4 (3.17)	5 (5.68)	12 (10.91)	7 (8.97)	5 (15.63)
		Association in the groups	*P*	OR	CI								
		JIA vs. C	<0.05	2.855	1.338—6.093								
		JIpA vs. C	**<0.001**	**5.230**	2.185—12.517								
4-LH-7	**A-G**-T-G	—	—	—	3	3	—	3	3	—	—	—	—
4-LH-8	**A-G**-T-A	—	—	—	2	2	—	2	2	—	—	—	—
4-LH-9	C-**G**-C-G	—	—	—	1	1	—	1	1	—	—	—	—
4-LH-10	C-**G**-C-A	—	—	—	2	2	—	2	2	—	—	—	—

Superscripts "Ref" and "R" indicate the reference and risk haplotypes respectively. Single nucleotide polymorphism loci in the 1—2—3—4 haplotypes are given in the rs2277460—rs2295826—rs2295827—rs2348071 sequence. Risk alleles are indicated in bold and underlined. Probability of association < 0.002 is indicated in bold. Frequencies of the 4-LH7, 4-LH8, 4-LH9 and 4-LH10 haplotypes rare/absent in all groups (<3%) are not indicated in the table. In statistical analysis the 4-LH1 haplotype was considered as reference.
4-LG = four-loci genotype; JIA = juvenile idiopathic arthritis; JIoA = juvenile idiopathic oligoarthritis; JIpA = juvenile idiopathic polyarthritis; C = control; F = female; M = male.

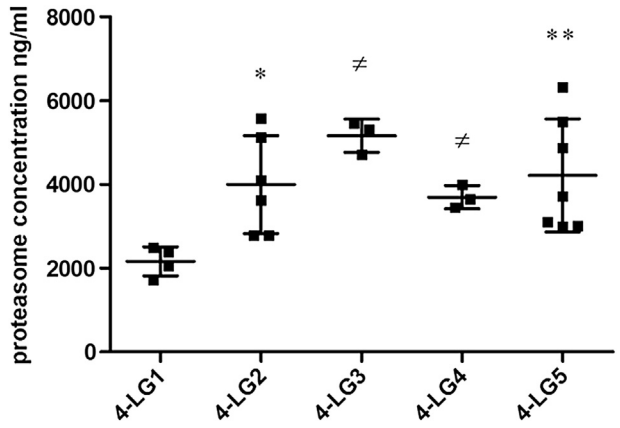


Figure 1 Plasma proteasome level in juvenile idiopathic oligoarthritis patients. Carriers of different four-locus genotypes (4-LGs), 4-LG1, 4-LG2, 4-LG3, 4-LG4, and 4-LG5, were represented by four, six, three, three, and seven patients respectively. * $p < 0.05$, ** $p < 0.001$, # $p > 0.05$.

having the only risk allele at the rs2277460 locus showed strong ($p < 0.001$) association with JIpA.

3.4. Genotype dependent plasma proteasome levels in JIoA patients

Plasma samples were available for 23 JIoA females, carriers of five different 4-LGs (Figure 1). Females of 4-LG1 exhibited a plasma proteasome concentration of approximately 2000 ng/mL, which is similar to previously reported plasma proteasome levels for healthy donors.[25–27] Careers of the 4-LG2 and 4-LG5 genotypes exhibited significantly ($p < 0.05$ and $p < 0.001$) higher plasma proteasome levels. High plasma proteasome levels were also detected in females of 4-LG3; however, the small number of patients in this group did not allow the results to reach statistical significance.

3.5. Eventual functional significance of the SNPs allelic variants

Figure 2 summarizes results of the *in silico* analysis of the functional significance of allele substitutions (only loci detected as JIA susceptible were taken into account) evaluated on the eventual sequence affinity to TFs and splicing signals similarity, and on the homology to known microRNAs and their precursors.

The major allele of only the rs2295826 locus potentially assists in sequence affinity to TFs. These are proteins of the CREB, MYT1 and PARF families. The rs2295826 minor allele appears to abolish any sequence affinity to TFs. Minor alleles at the rs2277460, rs2295827, and rs2348071 loci potentially assist in the binding of proteins belonging to the BARBIE box, CART, BRN5, LHXF, HOXF, HBOX, and MEF2 families.

Major alleles of both the rs2295826 and rs2295827 loci potentially assist in the generation of additional branch points; the rs2295826 major allele creates a splice site acceptor and targets for the hnRNP A1; the hnRNP A1 target motif is also generated in presence of both the rs2277460 and rs2348071 minor alleles. Major alleles of the rs2295826 and rs2348071 and minor alleles of all loci besides the rs2295826 could potentially change the sequence similarity to a number of splicing enhancers and/or silencers (Figure 2 and Suppl Table 3).

The rs2295827 major and rs2295826 minor allele increase sequence similarity to hsa-miR-603 and hsa-miR-5584-3p, respectively (Figure 2).

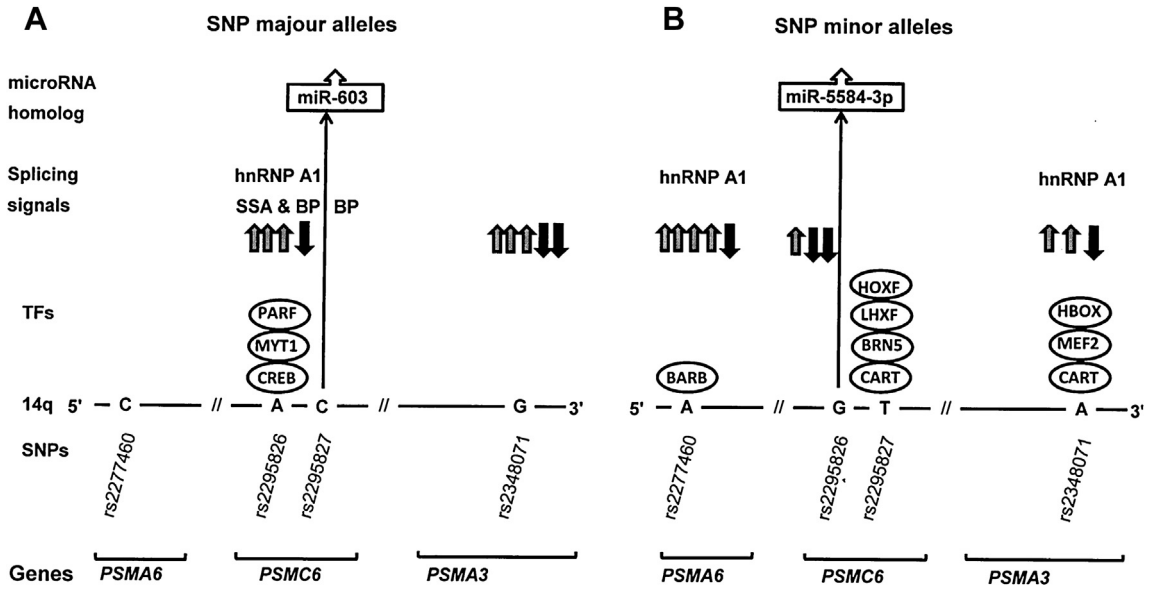


Figure 2 Consequences of the nucleotide substitutions on functional potential of the genome site harboring juvenile idiopathic arthritis associated single nucleotide polymorphisms of the *PSMA6*, *PSMC6*, and *PSMA3* genes. Only family names are indicated for transcription factors. Upward and downward arrows indicate splicing enhancers and silencers, respectively. BP = branch point; SSA = splicing site acceptor. Details are given in Supplementary Table 3.

4. Discussion

The aim of the current study was to evaluate six SNPs belonging to the *PSMA6* (rs2277460 and rs1048990), *PSMA3* (rs2348071), *PSMB5* (rs11543947), and *PSMC6* (rs2295826 and rs2295827) proteasomal genes for association with JIA with adjustment by JIA subtype and sex.

From all SNPs analyzed, the rs1048990 (*PSMA6* c.-8C>G) located in the 5′- untranslated region of the gene is the most studied SNP of proteasomal genes, which has been widely genotyped for association with cardiovascular diseases,[Suppl 1–12,14,15] T2DM,[Suppl 3,14] and children obesity.[Suppl 16] Summarizing the results obtained by different teams, we suggested a potential of the rs1048990 to influence JIA susceptibility. However, we did not find any association between the rs1048990 polymorphism and JIA in Latvians. Similarly, locus did not show any association with obesity in Latvian children.[Suppl 16] In turn, the rs2277460 of the promoter region of the same gene (*PSMA6* c.-110C>A) has been detected as JIA susceptible locus. This conclusion is based on the results of the subtype- and sex-specific disease association with rare allele A, heterozygous SLG, 4-LG6 heterozygous at the rs2277460, and 4-LH6 haplotype both having the rs2277460 rare allele in its structure (Tables 1–3).

It had been reported that the rs2230087 polymorphism of the *PSMB5* gene (3′-untranslated region) is associated with T2DM.[Suppl 9] In our study we were interested in the rs11543947 of the same gene locating in exon 1 or intron 1 (*PSMB5* c.70C>T or *PSMB5* c.-112+300C>T, respectively) depending of transcript variant. This SNP did not show any association with JIA in our study.

The rs2295826 and rs2295827 loci locate in close vicinity to each other in intron 1 of the *PSMC6* gene (*PSMC6* c.86-104A>G and *PSMC6* c.86-46C>T respectively). Linkage between loci is not complete, and rare GC haplotypes were observed in our study similar to data obtained for Tuscans in Italy. Rare alleles of these loci and their risk SLGs showed (dominant model) JIoA subtype-specific association by themselves and as component of risk 4-LG5 genotype and risk 4-LH4 haplotype.

The rs2348071 locus belongs to intron 7 of the *PSMA3* gene (PSMA3 c.543+138G>A or c.522+138G>A depending of transcript variant). Heterozygous genotype GA was found to be associated mostly with JIpA. In 4-LG structures, the rs2348071 heterozygotes were involved in both risk 4-LG5 and 4-LG6. Interestingly, the rs2348071 heterozygotes were implicated previously as an obesity risk factor in Latvian children with a family history of obesity.[Suppl 16]

It is important to note that strength of association with the disease was much stronger for combination of several risk SLGs than for any individual risk SLG. Therefore, we have reported here for the first time evidence of an association between JIA and genetic variants in the *PSMA6/PSMC6/PSMA3* gene cluster represented by combinations of at least two risk SLGs in a particular 4-LG, namely 4-LG5 (risk rs2295826/rs2295827/rs2348071) and 4-LG6 (risk rs2277460/rs2348071). The 4-LG1 having no risk SLGs in its composition showed strong association with healthy phenotype ($p < 0.001$).

The JIA-associated SNPs discovered in our study potentially could be themselves primarily susceptible to disease or linked with other primary genetic variations linked to disease. It appears that both scenarios are possible. Concerning chromosome 14, several loci potentially susceptible to autoimmune diseases have been reported in different human populations.[17–24] The functional significance of the discovered allele substitutions is to be clarified. We attempted to shed more light on the problem using two approaches.

First, we evaluated plasma proteasome level in JIoA females of different 4-LGs and found significantly increased levels of plasma proteasomes in JIoA female carriers of the 4-LG5 risk genotype in comparison to carriers of protective 4-LG1 ($p < 0.001$) genotype. Earlier, circulating proteasomes were suggested as markers in autoimmune diseases.[8] Concentration of circulating proteasomes was shown to be substantially elevated in patients with rheumatoid[8] and psoriatic[9] arthritis. The 20S proteasome has been identified as a target of the humoral autoreactive immune response in patients with systemic inflammatory diseases including autoimmune myositis,[10] primary Sjögren's syndrome,[11] dilated cardiomyopathy,[12] systemic lupus erythematosus,[10,13,14] multiple sclerosis,[15] and psoriatic arthritis.[14] The proteasomal inhibitor MG132 has been reported to reduce the severity of arthritis and reverse the pain behavior in the arthritic rat models.[16] To our knowledge, plasma levels of factors within the UPS have not been yet evaluated in JIA, and here we report data on that for the first time.

Second, we have evaluated eventual functional significance of allele substitutions on sequence affinity to TFs, splicing signals similarity and on the homology to known microRNAs and their precursors.

The major allele of the rs2295826 potentially assists to sequence affinity for TFs of CREB, MYT1, and PARF families known to be involved in regulation of multiple physiological processes and control of the circadian clock.[28–30] CREB-related TFs are especially interesting with respect of JIA pathogenesis, as they are known to be essential for osteoblast differentiation and function,[28] and they have been implicated in immune response.[29] It is of interest that expression of CREB, MYT1, and PARF proteins potentially could share the same epigenetic mechanism of regulation by hsa-miR-1264 originated from the X chromosome and potentially be differently expressed and differently involved in epigenetic network in females and males (data not shown).

The presence of a minor allele at the rs2277460 locus creates a binding site to the BARBIE box proteins reported to be involved in signal transduction pathways during development[31] and modulation of innate immunity.[32] Sequences having minor alleles at the rs2295827 and rs2348071 sites can potentially bind CART proteins responsible for bone and cartilage development.[33] Moreover, the rs2295827 and rs2348071 minor alleles could assist in sequence affinity to BRN5, LHXF, MEF2, and HBOX factors known to mediate transcriptional control of neuronal differentiation[34–37] and HOXF family NANOG.01 factor generally involved in signal transduction pathways during development.[38]

Similar to TFBSs, patterns of predicted splicing signals are allele specific. The rs2295826 and rs2348071 loci create a number of allele-specific targets for splicing enhancers and silencers. Only minor allele of the rs2277460 and only major allele of the rs2295827 is functional in this respect. Nucleotide substitutions at the rs2277460, rs2295826, and

rs2348071 define affinity of corresponding sequences to the hnRNP A1 known as alternative splicing repressor[39] and factor facilitating processing of specific microRNAs.[40] The above mentioned allele-specific differences in spectra of splicing signals could potentially significantly affect genes splicing activity and affect UPS efficiency.

Therefore, all of the above types of associations revealed and data on functional significance of allele substitutions are in good agreement between themselves and provide evidence that: (1) variations at the rs2277460, rs2295826, rs2295827, and rs2348071 loci could assist JIA susceptibility; (2) combination of the rs2348071 and rs2295826 and/or rs2295827 risk genotypes (4-LG5) represents the genetic module highly associated with both JIoA and JIpA and JIA female phenotype and plasma proteasome level in JIoA females; (3) combination of the rs2348071 and rs2277460 risk genotypes (4-LG6) represents the genetic module presumably associated with JIpA and male phenotype; (4) nucleotide substitutions affect the potential of encompassing sequences to create splicing signals, TFBSs and microRNAs; and (5) the *PSMA6*/*PSMC6*/*PSMA3* genetic variants and multiloci genetic modules could be suggested as JIA subtype- and sex-specific risk factors.

In conclusion it should be mentioned that, despite the rather small number (174/191 of cases/controls), this study can be considered as representative for the small Latvian population (< 2 million). Keeping in mind that JIA unites several clinically different subtypes, and that this disease tends to affect females more than males, we have applied stratification by JIA-subtype and sex. Due to the small subgroups, we sometimes could not reach significance. However, when significance was achieved, we obtained interesting results to be investigated with reference to other populations.

Conflicts of interest

All contributing authors declare no conflicts of interest.

Acknowledgments

Support for this work was partly provided by the Latvian National Research Program 2014 and European Social Foundation Project No 2013/0043/1DP/1.1.1.2.0/13/APIA/VIAA/002.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.pedneo.2014.01.007.

References

1. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet* 2007;**369**:767–78.
2. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004;**31**:390–2.
3. Prahalad S, Glass DN. A comprehensive review of the genetics of juvenile idiopathic arthritis. *Pediatr Rheumatol Online J* 2008;**6**:11.
4. Thompson SD, Barnes MG, Griffin TA, Grom AA, Glass DN. Heterogeneity in juvenile idiopathic arthritis: impact of molecular profiling based on DNA polymorphism and gene expression patterns. *Arthritis Rheum* 2010;**62**:2611–5.
5. Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, Manolio T, et al. Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet* 2011;**89**:607–18.
6. Kloetzel PM. Antigen processing by the proteasome. *Nat Rev Mol Cell Biol* 2001;**2**:179–87.
7. Sutoh Y, Kondo M, Ohta Y, Ota T, Tomaru U, Flajnik MF, et al. Comparative genomic analysis of the proteasome β5t subunit gene: implications for the origin and evolution of thymoproteasomes. *Immunogenetics* 2012;**64**:49–58.
8. Egerer K, Kuckelkorn U, Rudolph PE, Rückert JC, Dörner T, Burmester GR, et al. Circulating proteasomes are markers of cell damage and immunologic activity in autoimmune diseases. *J Rheumatol* 2002;**29**:2045–52.
9. Henry L, Le Gallic L, Garcin G, Coux O, Jumez N, Roger P, et al. Proteolytic activity and expression of the 20S proteasome are increased in psoriasis lesional skin. *Br J Dermatol* 2011;**165**:311–20.
10. Feist E, Dörner T, Kuckelkorn U, Schmidtke G, Micheel B, Hiepe F, et al. Proteasome alpha-type subunit C9 is a primary target of autoantibodies in sera of patients with myositis and systemic lupus erythematosus. *J Exp Med* 1996;**184**:1313–8.
11. Feist E, Kuckelkorn U, Dörner T, Dönitz H, Scheffler S, Hiepe F, et al. Autoantibodies in primary Sjögren's syndrome are directed against proteasomal subunits of the alpha and beta type. *Arthritis Rheum* 1999;**42**:697–702.
12. Voigt A, Bartel K, Egerer K, Trimpert C, Feist E, Gericke C, et al. Humoral anti-proteasomal autoimmunity in dilated cardiomyopathy. *Basic Res Cardiol* 2010;**105**:9–18.
13. Arribas J, Luz Rodríguez M, Alvarez-Do Forno R, Castaño JG. Autoantibodies against the multicatalytic proteinase in patients with systemic lupus erythematosus. *J Exp Med* 1991;**173**:423–7.
14. Colmegna I, Sainz Jr B, Citera G, Maldonado-Cocco JA, Garry RF, Espinoza LR. Anti-20S proteasome antibodies in psoriatic arthritis. *J Rheumatol* 2008;**35**:674–6.
15. Fissolo N, Kraus M, Reich M, Ayturan M, Overkleeft H, Driessen C, et al. Dual inhibition of proteasomal and lysosomal proteolysis ameliorates autoimmune central nervous system inflammation. *Eur J Immunol* 2008;**38**:2401–11.
16. Ahmed AS, Li J, Ahmed M, Hua L, Yakovleva T, Ossipov MH, et al. Attenuation of pain and inflammation in adjuvant-induced arthritis by the proteasome inhibitor MG132. *Arthritis Rheum* 2010;**62**:2160–9.
17. Arya R, Hare E, Del Rincon I, Jenkinson CP, Duggirala R, Almasy L, et al. Effects of covariates and interactions on a genome-wide association analysis of rheumatoid arthritis. *BMC Proc* 2009;**3**:S84.
18. Burgner D, Davila S, Breunis WB, Ng SB, Li Y, Bonnard C, et al. A genome-wide association study identifies novel and functionally related susceptibility loci for Kawasaki disease. *PLoS Genet* 2009;**5**:e1000319.
19. Chistyakov DA, Savost'anov KV, Turakulov RI, Nosikov VV. Genetic determinants of Graves disease. *Mol Genet Metab* 2000;**71**:66–9.
20. Cornélis F, Fauré S, Martinez M, Prud'homme JF, Fritz P, Dib C, et al. New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. *Proc Natl Acad Sci U S A* 1998;**95**:10746–50.
21. Stuart PE, Nair RP, Ellinghaus E, Ding J, Tejasvi T, Gudjonsson JE, et al. Genome-wide association analysis identifies three psoriasis susceptibility loci. *Nat Genet* 2010;**42**:1000–4.

22. Tomer Y, Davies TF. The genetic susceptibility to Graves' disease. *Baillieres Clin Endocrinol Metab* 1997;**11**:431—50.
23. Sjakste T, Eglite J, Sochnevs A, Marga M, Pirags V, Collan Y, et al. Microsatellite genotyping of chromosome 14q13.2-14q13 in the vicinity of proteasomal gene PSMA6 and association with Graves' disease in the Latvian population. *Immunogenetics* 2004;**56**:238—43.
24. Sjakste T, Trapina I, Rumba-Rozenfelde I, Lunin R, Sugoka O,Sjakste N. 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* 2010;**29**:543—51.
25. Lavabre-Bertrand T, Henry L, Carillo S, Guiraud I, Ouali A, Dutaud D, et al. Plasma proteasome level is a potential marker in patients with solid tumors and hemopoietic malignancies. *Cancer* 2001;**92**:2493—500.
26. Stoebner PE, Lavabre-Bertrand T, Henry L, Guiraud I, Carillo S, Dandurand M, et al. High plasma proteasome levels are detected in patients with metastatic malignant melanoma. *Br J Dermatol* 2005;**152**:948—53.
27. Minagar A, Ma W, Zhang X, Wang X, Zhang K, Alexander JS, et al. Plasma ubiquitin-proteasome system profile in patients with multiple sclerosis: correlation with clinical features, neuroimaging, and treatment with interferon-beta-1b. *Neurol Res* 2012;**34**:611—8.
28. Wang Q, Maillard M, Schibler U, Burnier M, Gachon F. 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* 2010;**299**:R1013—9.
29. Male V, Nisoli I, Gascoyne DM, Brady HJ. E4BP4: an unexpected player in the immune response. *Trends Immunol* 2012;**33**:98—102.
30. Elefteriou F, Ahn JD, Takeda S, Starbuck M, Yang X, Liu X, et al. Leptin regulation of bone resorption by the sympathetic nervous system and CART. *Nature* 2005;**434**:514—20.
31. Dozmorov M, Wu W, Chakrabarty K, Booth JL, Hurst RE, Coggeshall KM, et al. 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* 2009;**9**:152.
32. Arbouzova NI, Bach EA, Zeidler MP. Ken & barbie selectively regulates the expression of a subset of Jak/STAT pathway target genes. *Curr Biol* 2006;**16**:80—8.
33. Furukawa K, Iioka T, Morishita M, Yamaguchi A, Shindo H, Namba H, et al. Functional domains of paired-like homeoprotein Cart1 and the relationship between dimerization and transcription activity. *Genes Cells* 2002;**7**:1135—47.
34. Gill GN. Decoding the LIM development code. *Trans Am Clin Climatol Assoc* 2003;**114**:179—89.
35. Phillips K, Luisi B. The virtuoso of versatility: POU proteins that flex to fit. *J Mol Biol* 2000;**302**:1023—39.
36. She H, Mao Z. Regulation of myocyte enhancer factor-2 transcription factors by neurotoxins. *Neurotoxicology* 2011;**32**:563—6.
37. Uzumcu A, Karaman B, Toksoy G, Uyguner ZO, Candan S, Eris H, et al. 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* 2009;**52**:315—20.
38. Ho B, Olson G, Figel S, Gelman I, Cance WG, Golubovskaya VM. Nanog increases focal adhesion kinase (FAK) promoter activity and expression and directly binds to FAK protein to be phosphorylated. *J Biol Chem* 2012;**287**:18656—73.
39. Clower CV, Chatterjee D, Wang Z, Cantley LC, Vander Heiden MG, Krainer AR. The alternative splicing repressors hnRNP A1/A2 and PTB influence pyruvate kinase isoform expression and cell metabolism. *Proc Natl Acad Sci U S A* 2010;**107**:1894—9.
40. Michlewski G, Guil S, Cáceres JF. Stimulation of pri-miR-18a processing by hnRNP A1. *Adv Exp Med Biol* 2010;**700**:28—35.

PUBLICATION III

3.3. Genetic variants in the *PSMA6*, *PSMC6* and *PSMA3* genes associated with childhood asthma in Latvian and Taiwanese populations

ISSN 0233–7657. Biopolymers and Cell. 2014. Vol. 30. N 5. P. 377–387 doi: http://dx.doi.org/10.7124/bc.0008B5

BIOMEDICINE

UDC 577.21 + 616.248

Genetic variants in the *PSMA6*, *PSMC6* and *PSMA3* genes associated with childhood asthma in Latvian and Taiwanese populations

N. Paramonova[1^], L. Shi-Shin Wu[2^], I. Rumba-Rozenfelde[3, 4], J.-Y. Wang[5], N. Sjakste[1, 6], T. Sjakste[1]

[1]Genomics and Bioinformatics, Institute of Biology, University of Latvia
3, Miera Str., Salaspils, Latvia, LV-2169

[2]Institute of Biomedicine, Tzu-Chi University
701, Zhongyang Rd., Sec 3., Hualien 97004, Taiwan

[3]Faculty of Medicine, University of Latvia
1a, Sharlotes Str., Riga, Latvia, LV-1001

[4]University Children Hospital
20, Juglas Str., Riga, Latvia, LV-1079

[5]Division of Allergy and Clinical Immunology, College of Medicine, National Cheng Kung University
1, University Road, Tainan City 701, Tainan, Taiwan

[6]Latvian Institute of Organic Synthesis
21, Aizkraukles Str., Riga, Latvia, LV-1006

[^]Equal contribution

tanja@email.lubi.edu.lv

Proteasomes mediate functional realization of signaling proteins implicated in asthma pathogenesis. ***Aim****. To evaluate main and sex-specific association between the PSMA6, PSMC6 and PSMA3 proteasomal genes variations and childhood asthma in Latvians and Taiwanese.* ***Methods****. SNPs rs2277460, rs1048990, rs2295826, rs2295827 and rs2348071 were genotyped in 102 Latvian and 159 Taiwanese cases for comparison with genetic diversity in populations (191 and 1097 subjects respectively).* ***Results****. Haplotype CGACG showed strong ($P < 0.0001$) association with asthma risk in both populations. All loci heterozygous genotypes and haplotype CCGTA were identified as asthma risk factors in Latvians; rs1048990 and rs2348071 GG homozygotes and rs2295826 and rs2295827 heterozygotes showed asthma risk and protective effect in Taiwanese females respectively. The multi locus genotypes homozygous for alleles being common in Latvian population were identified as protective in Latvians and disease susceptible in Taiwanese.* ***Conclusions****. Our results suggest an association of the 14q13-23 proteasomal genes polymorphisms with the childhood asthma in Latvians and Taiwanese and highlight risk and/or protective factors being the same or different between the populations.*

Keywords: chromosome 14q13-23, SNPs, PSMA6, PSMC6, PSMA3, childhood asthma.

Introduction. Asthma is a chronic inflammatory disease caused by complex gene–gene and gene–environment interactions with hyper-responsiveness to various nonspecific stimuli [1–3] being to a large extent genetically heterogeneous between human populations [4]. A number of genes implicated in asthma encode various signaling proteins and transcription factors including those driven by NF-κB signaling pathways [5–7].

However, interplay of multiple risk alleles and/or genotypes and primary driver of the disease remains still unclear.

In eukaryotes, processing and degradation of vast majority of regulatory proteins are mediated by ubiquitin-proteasome system (UPS). Proteasomes, key UPS enzymatic complex possess several types of peptidase, endoribonuclease, protein-chaperone and DNA-helicase activities [8–10] allowing strict control and coordination of all steps of gene expression, genes and proteins networks and processes of genome–environment interaction. Insufficient proteasome function was implicated in pathophysiology of various acute and chronic lung diseases and their complications [11–14] and potentially could be a consequence of particular proteasomal genes structural variations.

Multiple studies including several GWAS analyses, indicated the 14q11-24 genome region as susceptible to asthma [15–21]. This genomic region possesses a cluster of proteasomal genes including the *PSMA6*, *PSMC6* and *PSMA3* genes implicated previously in susceptibility to autoimmunity [22–24], type 2 diabetes mellitus [25, 26], cardio-vascular disorders [27] and population adaptation to environment [28]. It appears that there is a large potential for the 14q proteasomal genes association studies to provide novel insights into the bronchial asthma (BA) pathogenesis in particular human populations and in general.

Aim of the current study was to genotype five single nucleotide polymorphisms (SNPs) belonging to the *PSMA6* (rs2277460 and rs1048990), *PSMA3* (rs2348071), and *PSMC6* (rs2295826 and rs2295827) proteasomal genes and evaluate main and sex-specific association between variations of these genes and asthma in Latvians and Taiwanese.

Materials and methods. One hundred two children (28 girls) aged under five and 159 (69 girls) aged under three represented Latvian (LV) and Taiwanese (TW) asthma groups respectively. LV BA patients were enrolled from the outpatient clinic of P. Stradins Clinical University Hospital and Children Clinical University Hospital «Gailezers» in Riga, Latvia.

TW study subjects were enrolled from elementary school for allergy diseases screen Taoyuan General Hospital, Taiwan. All patients were diagnosed with mild or moderate persistent asthma according to the guidelines of the Global Initiative for Asthma (GINA; http://www.ginasthma.org/local/uploads/files/GINA_Under5_Pocket_20091_1.pdf). The studies were approved by the Central Medical Ethics Commission of the Republic of Latvia Ministry of Health and Ethics Committee of Taoyuan General Hospital, Taiwan. Informed consent was obtained from parents of study participants.

Latvian and Taiwanese control groups of 191 (age = = 54.8 ± 18.6; 117 women) and 1097 (aged under five; 558 girls) participants respectively, were described and genetic diversity of SNPs of interest was studied previously [28] providing primary genotyping data to be used in current study to evaluate asthma main effects for each particular SNP, construct multi locus genotypes and stratify controls by sex to reveal asthma sex specific associations in single- and multi-locus models.

DNA extraction and genotyping technologies were the same as in [28].

For quality control, of the 16 randomly chosen samples per each marker were genotyped in duplicate in different experiments for asthma samples from both Latvian and Taiwanese collections. The concordance of the genotyping was 100 %. The chromosome 14 GRCh37.p5 assembly (NCBI reference sequence: NC_000014.8) sequence information was used for loci description.

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, single locus genotypes (SLGs) and alleles frequencies were estimated by direct gene counting. DnaSP version 5 (http://www.ub.es/dnasp/ [29] was used to reconstruct the haplotypes from un-phased genotypes, evaluate the nucleotide and haplotype genetic diversity and pairwise linkage disequilibrium (LD) between the loci (D' and r^2). Both the two-tailed Fisher's exact test and the χ^2 test were applied to evaluate the linkage between the rs2295826 and rs2295827 polymorphic sites at three p-value levels ($p < 0.05$; $p < 0.01$; $p < 0.001$). The Bonferroni correction included in DnaSP analysis was taken into account to support the significance of the revealed disequilibrium ($\alpha' = = 0.05$).

Deviation from the Hardy-Weinberg equilibrium and differences between case and control groups in al-

lele, genotype and haplotype frequencies as well as permutation test (Monte Carlo method/number of simulations = 10000) were evaluated by χ^2 using XLSTAT 2013 software for Windows. Dominant, recessive, over dominant and multiplicative genetic models for every individual locus were designed according to Lewis [30] and analysed by using 2 × 2 contingency tables. Odds ratio (OR) more than 2 and less than 0.5 was considered to be clinically significant. Stratification was performed by sex.

Results and discussion. In both Latvian and Taiwanese sample collections the genotyping call rate was 100 % for all markers; alleles and genotypes frequencies are given in Tables 1 and 2. The rs2348071 being in Latvian patients in HWE, significantly ($P < 0.001$) deviated from equilibrium in Taiwanese. Other markers were found to be in HWE in both LV and TW patients. The rs2295826 and rs2295827 were observed in complete ($D' = 1$, $r^2 = 1$) and slightly disrupted ($D' = 1$, $r^2 = 0.896$) LD in Latvians and Taiwanese respectively.

The distributions of alleles and genotypes in case groups were compared with those previously identified in the populations [28] and the data on single-locus association are summarized in Table 1 and Table 2. In Latvians all five loci showed the asthma main effect for rare alleles and heterozygous genotypes. The rs1048990 was associated with the disease in both females and males. The asthma susceptibility of resting loci was characterized by nonadditivity; that was, the rs2277460 and rs2348071 were associated with asthma in females, and the rs2295826 and rs2295827 were associated with asthma in males. In Taiwanese, the rs2277460 appears to be asthma neutral and resting loci showed the disease susceptibility only in females. The asthma risk effect was observed for the rs1048990 and rs2348071 GG genotypes and rs2348071 allele G. Rare alleles and heterozygous genotypes of the rs2295826 and rs2295827 showed female-specific asthma protective effect.

The multi-locus genotypes showing asthma risk or protective effect in any of our populations are listed in Table 3 (see suppl.). In Latvians a statistically significant protective effect was observed for all variants of multi locus genotypes being homozygous for the alleles common in the population. The 5LG of CC/CC/AA/CC/GG configuration being protective in Latvians (OR = = 0.322 [0.196 – 0.651]), showed the asthma risk effect in Taiwanese females (OR = 2.911 [1.327–6.387]). Similarly, being protective in Latvians the AA/CC/GG (rs2295826/rs2295827/rs2348071) genotype appears to be the disease susceptible in Taiwanese. In Latvians, the risk effect was observed for the rs2277460/rs2348071, rs1048990/rs2348071, rs1048990/rs2295826/rs2295827 and rs2295826/rs2295827/rs2348071 genotypes being simultaneously heterozygous at all loci involved. All mentioned genotypes were neutral in Taiwanese. In contrast, the rs2295826/rs2295827/rs2348071 genotype of AG/CT/AA configuration being neutral in Latvians, showed the protective effect in Taiwanese females.

The data of haplotype analysis are given in Table 4 (see suppl.). Haplotype diversity was higher in Latvian patients than in the population and did not differ between the cases and population in Taiwanese. The most frequent in Latvians the Hap1 (CCACG) showed male specific asthma protective (OR = 0.633 [0.416–0.963]) effect in this population. The Hap6 (CGACG) showed strong ($P < 0.0001$) association with the asthma risk in both Latvians (OR = 4.525 [2.286–8.958]) and Taiwanese (OR = 2.448 [1.763–3.399]) for both females and males. Minor in both populations and Taiwanese cases the Hap9 (CCGTA) was strongly ($P < 0.0001$) associated with the asthma phenotype in Latvians.

Identification of the genetic risk factors for asthma is complicated by potential interaction of genes and metabolic pathways, genotype with sex and environment; most of the reported asthma genes were not replicated across populations [4, 31]. The 14q11-24 chromosomal region is one of well replicated asthma susceptibility loci [15–21]; the *PSMA6*, *PSMA3* and *PSMC6* proteasomal genes located in the region were implicated earlier in susceptibility to autoimmunity [22–24], inflammation [25–27] and historical and geographical adaptation [28].

In this paper we provide for the first time the evidence that polymorphism in the *PSMA6*, *PSMC6*, and *PSMA3* proteasomal genes may contribute to the risk of childhood asthma in both Latvian and Taiwanese populations. The most remarkable finding of our study is that haplotype CGACG was revealed to be a strong ($P < 0.0001$) asthma risk factor in both Latvians and Taiwanese. Other identified asthma risk and protective single- and multi-locus genetic variants are different between two populations. The difference between human

Table 1
SNP allele and genotype distribution and data on association with paediatric asthma in Latvian population

Marker Allele or genotype	Distribution of alleles and genotypes, *n* (%)					
	BA patients			Controls*		
	Total (*n* = 102)	Females (*n* = 28)	Males (*n* = 74)	Total (*n* = 191)	Females (*n* = 117)	Males (*n* = 74)
rs2277460						
C	176 (86.27)	47 (83.93)	129 (87.16)	357 (93.46)	232 (94.87)	135 (91.22)
A	28 (13.73)	9 (16.07)	19 (12.84)	25 (6.54)	12 (5.13)	13 (8.78)
CC	74 (72.55)	19 (67.86)	55 (74.32)	166 (86.91)	105 (89.74)	61 (82.43)
CA	28 (27.45)	9 (32.14)	19 (25.68)	25 (13.09)	12 (10.26)	13 (17.57)
rs1048990						
C	165 (80.88)	44 (78.57)	121 (81.76)	348 (91.10)	211 (90.17)	137 (92.57)
G	39 (19.12)	12 (21.43)	27 (18.24)	34 (8.90)	23 (9.83)	11 (7.43)
	–	–	–	–	–	–
CC	65 (63.73)	16 (57.14)	49 (66.22)	158 (82.72)	95 (81.20)	63 (85.14)
CG	35 (34.31)	12 (42.86)	23 (31.08)	32 (16.75)	21 (17.95)	11 (14.86)
GG	2 (1.96)	–	2 (2.70)	1 (0.52)	1 (0.85)	–
rs2295826						
A	166 (81.37)	47 (83.93)	119 (80.41)	342 (89.53)	204 (87.18)	138 (93.24)
G	38 (18.63)	9 (16.07)	29 (19.59)	40 (10.47)	30 (12.82)	10 (6.76)
AA	66 (64.71)	19 (67.86)	47 (63.51)	155 (81.15)	90 (76.93)	65 (87.84)
AG	34 (33.33)	9 (32.14)	25 (33.78)	32 (16.75)	24 (20.51)	8 (10.81)
GG	2 (1.96)	–	2 (2.70)	4 (2.09)	3 (2.56)	1 (1.35)
rs2295827						
C	166 (81.37)	47 (83.93)	119 (80.41)	342 (89.53)	204 (87.18)	138 (93.24)
T	38 (18.63)	9 (16.07)	29 (19.59)	40 (10.47)	30 (12.82)	10 (6.76)
CC	66 (64.71)	19 (67.86)	47 (63.51)	155 (81.15)	90 (76.93)	65 (87.84)
CT	34 (33.33)	9 (32.14)	25 (33.78)	32 (16.75)	24 (20.51)	8 (10.81)
TT	2 (1.96)	–	2 (2.70)	4 (2.09)	3 (2.56)	1 (1.35)
rs2348071						
G	119 (58.33)	34 (60.71)	85 (57.43)	270 (70.68)	170 (72.65)	100 (67.57)
A	85 (41.67)	22 (39.29)	63 (42.57)	112 (29.32)	64 (27.35)	48 (32.43)
GG	35 (34.31)	8 (28.57)	27 (36.49)	102 (53.40)	65 (55.55)	37 (50.00)
GA	49 (48.04)	18 (64.29)	31 (41.89)	66 (34.56)	40 (34.19)	26 (35.14)
AA	18 (17.65)	2 (7.14)	16 (21.62)	23 (12.04)	12 (10.26)	11 (14.86)

*Data on an allele and genotype presentation in control group are given according to Sjakste *et al*. [24, 28]; *P* - probability calculated by χ^2 test;

Marker ID Allele or genotype	Statistics			
	Genetic model	Group	*P* (P_c)	OR [95 % CI]
rs2277460				
C	A *vs* C	Total	0.0038 (0.0045)	2.272 [1.293–3.933]
A	–	Females	0.0032 (0.0059)	3.702 [1.506–9.101]
CC	CA *vs* CC	Total	0.0023 (0.0029)	2.512 [1.379–4.578]
CA	–	Females	0.0031 (0.0066)	4.145 [1.568–10.957]
rs1048990				
C	G *vs* C	Total	0.0004 (0.0006)	2.419 [1.478–3.960]
G	–	Females	0.0167 (0.0238)	2.502 [1.157–5.342]
	–	Male	0.0054 (0.0085)	2.779 [1.339–5.768]
CC	CG *vs*	Total	0.0007 (0.0011)	2.596 [1.491–4.519]
CG	CC + GG	Females	0.0047 (0.0077)	3.429 [1.435–8.192]
GG	–	Male	0.0190 (0.0308)	2.583 [1.166–5.722]
rs2295826				
A	G *vs* A	Total	0.0056 (0.0064)	1.957 [1.213–3.158]
G	–	Male	0.0011 (0.0022)	3.363 [1.596–7.088]
AA	AG *vs*	Total	0.0012 (0.0016)	2.484 [1.424–4.334]
AG	AA + GG	Male	0.0008 (0.0008)	4.209 [1.783–9.937]
GG	–	–	–	–
rs2295827				
C	G *vs* A	Total	0.0056 (0.0064)	1.957 [1.213–3.158]
T	–	Male	0.0011 (0.0022)	3.363 [1.596–7.088]
CC	AG *vs*	Total	0.0012 (0.0016)	2.484 [1.424–4.334]
CT	AA + GG	Male	0.0008 (0.0008)	4.209 [1.783–9.937]
TT	–	–	–	–
rs2348071				
G	A *vs* G	Total	0.0026 (0.0039)	1.722 [1.208–2.454]
A	–	–	–	–
GG	GA *vs*	Total	0.0244 (0.0343)	1.751 [1.075–2.851]
GA	GG + AA	Females	0.0035 (0.0059)	3.465 [1.485–8.085]
AA	–	–	–	–

P_c – corrected probability calculated by Monte Carlo method with 10000 simulations.

Table 2
SNP allele and genotype distribution and data on association with paediatric asthma in Taiwanese population

Marker Allele or genotype	Distribution of alleles and genotypes, *n* (%)					
	BA patients			Controls*		
	Total (*n* = 159)	Females (*n* = 69)	Males (*n* = 90)	Total (*n* = 1097)	Females (*n* = 558)	Males (*n* = 539)
rs2277460						
C	315 (99.06)	137 (99.28)	179 (99.44)	1081 (99.27)	547 (99.01)	534 (99.54)
A	2 (0.63)	1 (0.72)	1 (0.56)	16 (0.73)	11 (0.99)	5 (0.46)
CC	157 (98.74)	68 (98.55)	89 (98.89)	1081 (98.54)	547 (98.03)	534 (99.07)
CA	2 (1.26)	1 (1.45)	1 (1.11)	16 (1.46)	11 (1.97)	5 (0.93)
rs1048990						
C	209 (65.72)	85 (61.59)	124 (68.89)	1480 (67.46)	738 (66.13)	742 (68.83)
G	109 (34.28)	53 (38.41)	56 (31.11)	714 (32.54)	378 (33.87)	336 (31.17)
CC	66 (41.51)	28 (40.58)	38 (42.22)	462 (42.12)	228 (40.86)	234 (43.42)
CG	77 (48.43)	29 (42.03)	48 (53.33)	556 (50.68)	282 (50.54)	274 (50.83)
GG	16 (10.06)	12 (17.39)	4 (4.44)	79 (7.20)	48 (8.60)	31 (5.75)
rs2295826						
A	286 (89.94)	129 (93.48)	157 (87.22)	1855 (84.55)	948 (84.95)	907 (84.14)
G	32 (10.06)	9 (6.52)	23 (12.78)	339 (15.45)	168 (15.05)	171 (15.86)
AA	129 (81.13)	60 (86.96)	69 (76.67)	778 (70.92)	404 (72.40)	374 (69.39)
AG	28 (17.61)	9 (13.04)	19 (21.11)	299 (27.26)	140 (25.08)	159 (29.05)
GG	2 (1.26)	–	2 (2.22)	20 (1.82)	14 (2.51)	6 (1.11)
rs2295827						
C	289 (90.88)	129 (93.48)	160 (88.89)	1872 (85.32)	962 (86.20)	910 (84.42)
T	29 (9.12)	9 (6.52)	20 (11.11)	322 (14.68)	154 (13.80)	168 (15.58)
CC	130 (81.76)	60 (86.96)	70 (77.78)	775 (70.64)	404 (72.40)	371 (68.83)
CT	29 (18.24)	9 (13.04)	20 (22.22)	322 (29.35)	154 (27.60)	168 (31.17)
TT	–	–	–	–	–	–
rs2348071						
G	126 (39.62)	61 (44.20)	65 (36.11)	759 (34.59)	380 (34.05)	379 (35.16)
A	192 (60.38)	77 (55.80)	115 (63.89)	1435 (65.41)	736 (65.95)	699 (64.84)
GG	40 (25.16)	19 (27.54)	21 (23.33)	204 (18.60)	97 (17.39)	107 (19.85)
GA	46 (28.93)	23 (33.33)	23 (25.56)	351 (31.99)	186 (33.33)	165 (30.61)
AA	73 (45.91)	27 (39.13)	46 (51.11)	542 (49.41)	275 (49.28)	267 (49.54)

*Data on an allele and genotype presentation in control group are given according to Sjakste *et al*. [28]; *P* – probability calculated by χ^2 test;

Marker ID Allele or genotype	Statistics			
	Genetic model	Group	*P* (P_c)	OR [95 % CI]
rs2277460				
C	–	–	–	–
A	–	–	–	–
CC	–	–	–	–
CA	–	–	–	–
rs1048990				
C	–	–	–	–
G	–	–	–	–
CC	GG *vs*	Females	0.0192 (0.0311)	2.237 [1.135–4.410]
CG	CC + CG	–	–	–
GG	–	–	–	–
rs2295826				
A	G *vs* A	Total	0.0114 (0.0117)	0.612 [0.418–0.896]
G	–	Females	0.0066 (0.0094)	0.394 [0.200–0.776]
AA	AG *vs*	Total	0.0094 (0.0091)	0.570 [0.372–0.873]
AG	AA + GG	Females	0.0266 (0.0346)	0.488 [0.220–0.911]
GG	–	–	–	–
rs2295827				
C	T *vs* C	Total	0.0076 (0.0099)	0.583 [0.392–0.868]
T	–	Females	0.0165 (0.0215)	0.436 [0.221-0.861]
CC	CT *vs* CC	Total	0.0023 (0.0024)	0.523 [0.344–0.796]
CT	–	Females	0.0093 (0.0117)	0.394 [0.194–0.799]
TT	–	–	–	–
rs2348071				
G	C *vs* A	Females	0.0185 (0.0228)	1.534 [1.74–2.192]
A	–	–	–	–
GG	GG *vs*	Females	0.0404 (0.0315)	1.806 [1.025–3.182]
GA	GA + AA	–	–	–
AA	–	–	–	–

P_c – corrected probability calculated by Monte Carlo method with 10000 simulation.

populations in asthma genetics is a well-known phenomenon described for a number of asthma susceptible loci having mainly ethnos specific differences in genetic diversity [4]. The Latvian and Taiwanese populations significantly differ in genetic diversity of loci studied here [28]. This suggests involvement of these loci in the processes of evolutional and/or geographical adaptation to environment and a potential for allele substitutions to have different ethnic specific influence on the human health and population morbidity [28].

Several associations revealed in our study showed non-additivity between sexes. In Latvians the rs2277460 and rs2348071 were associated with asthma in females, and the rs2295826 and rs2295827 were disease susceptible in males; in Taiwanese all asthma susceptible loci were limited to females. Sex specific differences in incidence, prevalence, and severity are also well known features of asthma epidemiology. Sex-specific associations with the disease have been recently reported for SNPs of several genes-candidates including the *IFNG* [31], *IL17F* [32], *TSLP* [33], *VDR* [34], and *KCNB1* [35] genes. Our analysis of the BA main effect in Latvian population is a subject to some limitation as sexes were not equally presented in both BA and control groups. Although a significant asthma main effect was detected for all five loci studied, only the rs1048990 showed an additive effect that was an association in both females and males. The replication study in additional larger cohorts represented by sexes equally is required to validate the results found in the current study for Latvian population.

Due to the pleiotropic effect, a frequent phenomenon in human complex traits and diseases [36], some loci of susceptibility may be shared among many autoimmune and other immune-mediated diseases [37, 38]. Earlier the genetic pleiotropic effect has been reported for asthma and obesity [39, 40] and for asthma and juvenile rheumatoid arthritis [41, 42]. Similarly, SNPs associated with asthma in our current study, previously have been found to be susceptible in Latvians to other immune-mediated pathologies including juvenile idiopathic arthritis [24, 43], children obesity [44] and multiple sclerosis [45]. The rs1048990 was widely genotyped in many human populations and reported as an ethnic specific risk factor for inflammation within the cardio-vascular system [27, 28].

All loci we have studied here belong to the non-coding regions of corresponding genes and nucleotide substitutions potentially may influence the gene expression through allele specific targeting of different regulatory elements. Among the allele-specific targets described by Sjakste with co-authors earlier [24, 28], several sites showed affinity to transcription factors and splicing signals implicated previously in immunity, lung function and lung pathology. The targeting of these regulatory proteins may influence asthma pathogenesis and needs to be mentioned in respect of current study. The rs2277460 ancestral allele C, the major allele in human populations over the world, appears to be functionally neutral. Substitution to A generates a target to hnRNP A1, a multifunctional protein implicated in the association with multiple promoter sequences and modulation of a number of transcriptional events [46]. The hnRNP A1 has been shown to play a key role in many human pathologies including lung cancer and response to viral pathogens [46, 47]. It may influence protein-protein interactions including those with participation of NF-κB [46] playing in turn a significant role in the asthma development and progression [5–7]. Additionally, it is involved in crosstalk with ubiquitin proteasome system at different levels of NF-κB and other regulatory proteins signaling pathways [48]. Allele A also assists to sequence affinity to the BARBIE box proteins found to be involved in inflammatory response of alveolar macrophages [49]. Substitution C → G at the rs1048990 was shown to influence the gene expression *in vivo* and *in vitro* [27, 50] and significantly change the sequence capacity to bind a number of splicing signals and transcription factors [28]. Rare allele G generates binding sites for the multifunctional proteins of p53 and DMRT families implicated in the processes of climatic [51] and evolutional [52] adaptation. The targeting of these proteins potentially could be involved in the mechanisms of natural selection and ethnos specific susceptibility to inflammation [27, 28].

Common allele A of the rs2295826 (first intron of the *PSMC6* gene) generates the targets for the mentioned above hnRNP A1 regulatory protein and for the transcription factor of CREB family involved in transcriptional control of many pro-inflammatory genes [53, 54] and implicated in asthma pathogenesis [55], asthma phenotypes and response to therapy [56].

The rs2348071 SNP strongly discriminates Latvians having a major allele G (about 70 %) and Taiwanese having a major ancestral allele A (about 70 %). Previously we have suggested [28] that transition A → G happened in Caucasians about 15,000 years ago was supported by positive selection. This mutation eliminates potential targets for hnRNP A1 and the transcription factors of CART family shown to be an essential participant of signaling respiratory network [57] and the MEF2 family implicated in transcriptional switch between metabolism and immunity [58].

Summarizing mentioned results we suggest that the nucleotide substitutions we have studied may significantly modulate the transcription of related genes and gene network in response to the inflammation and other environmental stimuli and influence the asthma susceptibility.

Conclusions. Our findings provide an evidence that single- and multi locus variations in the 14q13-23 *PSMA6*/*PSMC6*/*PSMA3* proteasomal genes cluster are associated with childhood asthma in Latvian and Taiwanese populations and could play an important role in asthma and other immune-mediated pathologies in both Caucasians and Asians, as either the risk or protective ethnic- and sex-specific genetic factors.

Identification of genetic variants susceptible to asthma and other immune-mediated pathologies, both common and different across populations, is important in understanding pathogenesis and phenotype variability of these multifactorial diseases. It might be a subject of thorough investigation in the nearest future.

Funding. This work was suported by the Latvia-Lithuania-Taiwan Cooperation Project No V3 20100512 and Latvian National Research Programme 2010.10.-4/VPP4.

Генетичні варіанти генів *PSMA6*, *PSMC6* і *PSMA3*, асоційовані з бронхіальною астмою у дітей латвійської і тайваньської популяцій

Н. Парамонова, Л. С.-С. Ву, І. Румба-Розенфелде, Дж.-Я. Ванг, Н. Сьяксте, Т. Сьяксте

Резюме

Протеасоми опосередковують реалізацію функцій сигнальних білків, залучених до патогенезу бронхіальної астми. ***Мета****. Оцінити загальну і залежну від статі асоціацію варіацій протеасомних генів PSMA6, PSMC6 і PSMA3 з бронхіальною астмою у дітей із Латвії і Тайваня.* ***Методи****. Однонуклеотидні поліморфізми rs2277460, rs1048990, rs2295826, rs2295827 і rs2348071 генотиповано у 102 хворих з Латвії і 159 – з Тайваня. Для порівняння взято контрольні групи, які представляють генетичне різноманіття в популяціях: 191 латвійських і 1097 тайваньських зразків.* ***Результати****. Гаплотип CGACG виявився тісно (P < 0.0001) асоційованим з ризиком розвитку астми в обох популяціях. Гетерозиготні генотипи за всіма локусами і гаплотип CCGTA ідентифіковано як фактор ризику для розвитку астми у жителів Латвії. Гомозиготи GG по rs1048990 і rs2348071 пов'язані з ризиком, а гетерозиготи по rs2295826 і rs2295827 проявляють захисний ефект з-поміж тайваньських жінок. Багатолокусні генотипи, гомозиготні за розповсюдженими в Латвії алелями, виявилися захисними для жителів Латвії, але пов'язаними з ризиком захворювання серед тайваньців.* ***Висновки****. Наші результати вказують на асоціацію поліморфізмів протеасомних генів локусу 14q13-23 з бронхіальною астмою серед дітей у латвійській і тайваньській популяціях, асоціація може бути пов'язана як з ризиком захворювання, так і з захисним ефектом. За даною ознакою популяції можуть різнитися або не різнитися.*

Ключові слова: хромосома 14q13-23, однонуклеотидні поліморфізми, PSMA6, PSMC6, PSMA3, бронхіальна астма у дітей.

Генетические варианты генов *PSMA6*, *PSMC6* и *PSMA3*, ассоциированные с бронхиальной астмой у детей в латвийской и тайваньской популяциях

Н. Парамонова, Л. С.-С. Ву, И. Румба-Розенфелде, Дж.-Я. Ванг, Н. Сьяксте, Т. Сьяксте

Резюме

Протеасомы опосредуют реализацию функций сигнальных белков, вовлеченных в патогенез бронхиальной астмы. ***Цель****. Оценить общую и зависимую от пола ассоциацию вариаций протеасомных генов PSMA6, PSMC6 и PSMA3 с бронхиальной астмой у детей из Латвии и Тайваня.* ***Методы****. Однонуклеотидные полиморфизмы rs2277460, rs1048990, rs2295826, rs2295827 и rs2348071 генотипированы у 102 больных из Латвии и 159 – из Тайваня. Для сравнения взяты контрольные группы, представляющие генетическое разнообразие в популяциях: 191 латвийских и 1097 тайваньских образцов.* ***Результаты****. Гаплотип CGACG оказался тесно (P < 0.0001) ассоциированным с риском развития астмы в обеих популяциях. Гетерозиготные генотипы по всем локусам и гаплотип CCGTA идентифицированы как фактор риска для развития астмы у жителей Латвии. Гомозиготы GG по rs1048990 и rs2348071 связаны с риском, а гетерозиготы по rs2295826 и rs2295827 проявляют защитный эффект среди тайваньских женщин. Многолокусные генотипы, гомозиготные по распространенным в Латвии аллелям, оказались защитными для жителей Латвии, но связанными с риском заболевания среди тайваньцев.* ***Выводы****. Наши результаты указывают на ассоциацию полиморфизмов протеасомных генов локуса 14q13-23 с бронхиальной астмой среди детей в латвийской и тайваньской популяциях, ассоциация может быть связана как с риском заболевания, так и с защитным эффектом. По данному признаку популяции могут отличаться или не отличаться.*

Ключевые слова: хромосома 14q13-23, однонуклеотидные полиморфизмы, PSMA6, PSMC6, PSMA3, бронхиальная астма у детей.

REFERENCES

1. *Broide DH*. Molecular and cellular mechanisms of allergic disease. *J Allergy Clin Immunol*. 2001;**108**(2 Suppl):S65–71.
2. *Galli SJ, Tsai M, Piliponsky AM.* The development of allergic inflammation. *Nature*. 2008;**454**(7203):445–54.
3. *Wu LS, Sjakste T, Sakalauskas R, et al.* The burden of allergic asthma in children: a landscape comparison based on data from Lithuanian, Latvian, and Taiwanese populations. *Pediatr Neonatol*. 2012;**53**(5):276–82.
4. *Leung TF, Ko FW, Sy HY, et al.* Differences in asthma genetics between Chinese and other populations. *J Allergy Clin Immunol.* 2014;**133**(1):42–8.
5. *Birrell MA, Hardaker E, Wong S, et al.* Ikappa-B kinase-2 inhibitor blocks inflammation in human airway smooth muscle and a rat model of asthma. *Am J Respir Crit Care Med*. 2005;**172**(8):962–71.
6. *Edwards MR, Bartlett NW, Clarke D, et al.* Targeting the NF-kappaB pathway in asthma and chronic obstructive pulmonary disease. *Pharmacol Ther*. 2009;**121**(1):1–13.
7. *Gagliardo R, Chanez P, Mathieu M, et al.* Persistent activation of nuclear factor-kappaB signaling pathway in severe uncontrolled asthma. *Am J Respir Crit Care Med*. 2003;**168**(10):1190–8.
8. *Konstantinova IM, Tsimokha AS, Mittenberg AG.* Role of proteasomes in cellular regulation. *Int Rev Cell Mol Biol*. 2008; **267**:59–124.
9. *Wang X, Guerrero C, Kaiser P, Huang L*. Proteomics of proteasome complexes and ubiquitinated proteins. *Expert Rev Proteomics*. 2007;**4**(5):649–65.
10. *Yano M, Kanesaki Y, Koumoto Y, et al.* Chaperone activities of the 26S and 20S proteasome. *Curr Protein Pept Sci*. 2005;**6**(2): 197–203.
11. *Albright JM, Romero J, Saini V, et al.* Proteasomes in human bronchoalveolar lavage fluid after burn and inhalation injury. *J Burn Care Res*. 2009;**30**(6):948–56.
12. *Majetschak M, Sorell LT, Patricelli T, Seitz DH, Knoferl MW*. Detection and possible role of proteasomes in the bronchoalveolar space of the injured lung. *Physiol Res*. 2009;**58**(3):363–72.
13. *Sixt SU, Peters J*. Extracellular alveolar proteasome: possible role in lung injury and repair. *Proc Am Thorac Soc.* 2010;**7**(1):91–6.
14. *Zemeckiene Z, Vitkauskiene A, Sjakste T, Sitkauskiene B, Sakalauskas R*. Proteasomes and proteasomal gene polymorphism in association with inflammation and various diseases. *Medicina (Kaunas)*. 2013;**49**(5):207–13.
15. *Hakonarson H, Bjornsdottir US, Halapi E, et al.* A major susceptibility gene for asthma maps to chromosome 14q24. *Am J Hum Genet*. 2002;**71**(3):483–91.
16. *Malerba G, Patuzzo C, Trabetti E, et al.* Chromosome 14 linkage analysis and mutation study of 2 serpin genes in allergic asthmatic families. *J Allergy Clin Immunol*. 2001;**107**(4):654–8.
17. *Mansur AH, Bishop DT, Holgate ST, Markham AF, Morrison JF*. Linkage/association study of a locus modulating total serum IgE on chromosome 14q13-24 in families with asthma. *Thorax*. 2004;**59**(10):876–82.
18. *Munthe-Kaas MC, Gerritsen J, Carlsen KH, et al.* Eosinophil cationic protein (ECP) polymorphisms and association with asthma, s-ECP levels and related phenotypes. *Allergy*. 2007;**62** (4):429–36.
19. *Tulah AS, Beghe B, Barton SJ, Holloway JW, Sayers I*. Leukotriene B4 receptor locus gene characterisation and association studies in asthma. *BMC Med Genet*. 2012;**13**:110.
20. *Ungvari I, Hullam G, Antal P, et al.* Evaluation of a partial genome screening of two asthma susceptibility regions using bayesian network based bayesian multilevel analysis of relevance. *PLoS One*. 2012;**7**(3):e33573.
21. *Zhang Y, Moffatt MF, Cookson WO*. Genetic and genomic approaches to asthma: new insights for the origins. *Curr Opin Pulm Med*. 2012;**18**(1):6–13.
22. *Sjakste T, Eglite J, Sochnevs A, et al.* Microsatellite genotyping of chromosome 14q13.2-14q13 in the vicinity of proteasomal gene PSMA6 and association with Graves' disease in the Latvian population. *Immunogenetics*. 2004;**56**(4):238–43.
23. *Sjakste T, Trapina I, Rumba-Rozenfelde I, Lunin R, Sugoka O, Sjakste N*. 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.* 2010;**29**(9):543–51.
24. *Sjakste T, Paramonova N, Rumba-Rozenfelde I, Trapina I, Sugoka O, Sjakste N*. Juvenile idiopathic arthritis subtype- and sex-specific associations with genetic variants in the *PSMA6*/*PSMC6*/*PSMA3* gene cluster. *Pediatr Neonatol.* 2014;**55**(5): 393–403.
25. *Sjakste T, Kalis M, Poudziunas I, et al.* Association of microsatellite polymorphisms of the human 14q13.2 region with type 2 diabetes mellitus in Latvian and Finnish populations. *Ann Hum Genet.* 2007;**71**(Pt 6):772–6.
26. *Sjakste T, Poudziunas I, Ninio E, et al.* SNPs of PSMA6 gene – investigation of possible association with myocardial infarction and type 2 diabetes mellitus. *Genetika*. 2007;**43**(4):553–9.
27. *Wang H, Jiang M, Zhu H, et al.* Quantitative assessment of the influence of PSMA6 variant (rs1048990) on coronary artery disease risk. *Mol Biol Rep*. 2013;**40**(2):1035–41.
28. *Sjakste T, Paramonova N, Wu L, et al*. *PSMA6* (rs2277460, rs1048990), *PSMC6* (rs2295826, rs2295827) and *PSMA3* (rs2348071) genetic diversity in Latvians, Lithuanians and Taiwanese. Meta Gene. 2014; **2**:283–98.
29. *Rozas J.* DNA sequence polymorphism analysis using DnaSP. *Methods Mol Biol*. 2009;**537**:337–50.
30. *Lewis CM*. Genetic association studies: design, analysis and interpretation. *Brief Bioinform*. 2002;**3**(2):146–53.
31. *Loisel DA, Tan Z, Tisler CJ, et al.* IFNG genotype and sex interact to influence the risk of childhood asthma. *J Allergy Clin Immunol*. 2011;**128**(3):524–31.
32. *Qian F, Zhang Q, Zhou L, et al.* Association between polymorphisms in IL17F and male asthma in a Chinese population. *J Investig Allergol Clin Immunol*. 2012;**22**(4):257–63.
33. *Hunninghake GM, Soto-Quiros ME, Avila L, et al.* TSLP polymorphisms are associated with asthma in a sex-specific fashion. *Allergy*. 2010;**65**(12):1566–75.
34. *Raby BA, Lazarus R, Silverman EK, et al.* Association of vitamin D receptor gene polymorphisms with childhood and adult asthma. *Am J Respir Crit Care Med*. 2004;**170**(10):1057–65.
35. *Seibold MA, Wang B, Eng C, et al.* An african-specific functional polymorphism in KCNMB1 shows sex-specific association with asthma severity. *Hum Mol Genet*. 2008;**17**(17):2681–90.
36. *Sivakumaran S, Agakov F, Theodoratou E, et al.* Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet*. 2011;**89**(5):607–18.
37. *Lee SH, Park JS, Park CS*. The search for genetic variants and epigenetics related to asthma. *Allergy Asthma Immunol Res*. 2011;**3**(4):236–44.
38. *Thompson SD, Barnes MG, Griffin TA, Grom AA, Glass DN*. Heterogeneity in juvenile idiopathic arthritis: impact of molecular profiling based on DNA polymorphism and gene expression patterns. *Arthritis Rheum*. 2010;**62**(9):2611–5.
39. *Hallstrand TS, Fischer ME, Wurfel MM, Afari N, Buchwald D, Goldberg J.* Genetic pleiotropy between asthma and obesity in a

community-based sample of twins. *J Allergy Clin Immunol*. 2005;**116**(6):1235–41.
40. *Murphy A, Tantisira KG, Soto-Quiros ME, et al.* PRKCA: a positional candidate gene for body mass index and asthma. *Am J Hum Genet*. 2009;**85**(1):87–96.
41. *Lee SH, Lee EB, Shin ES, et al.* The interaction between allelic variants of CD86 and CD40LG: a common risk factor of allergic asthma and rheumatoid arthritis. *Allergy Asthma Immunol Res*. 2014;**6**(2):137–41.
42. *Ramirez-Bello J, Jimenez-Morales S, Espinosa-Rosales F, et al.* Juvenile rheumatoid arthritis and asthma, but not childhood-onset systemic lupus erythematosus are associated with FCRL3 polymorphisms in Mexicans. *Mol Immunol.* 2013;**53**(4):374–8.
43. *Trapina I, Rumba-Rozenfelde I, Sjakste N, Sokolovska J, Sugoka O, Sjakste T*. Association study of genetic variants in the 14q11–14q13 proteasomal genes cluster with juvenile idiopathic arthritis (JIA) in Latvian population. Proceedings of the Latvian Academy of Sciences. Section B. 2009; **63**(4–5):214–31.
44. *Kupca S, Sjakste T, Paramonova N, et al.* Association of obesity with proteasomal gene polymorphisms in children. *J Obes.* 2013; **2013**:638154.
45. *Kalnina J, Paramonova N, Sjakste N, Sjakste T.* Study of association between polymorphisms in the PSMB5 (rs11543947) and PSMA3 (rs2348071) genes and multiple sclerosis in Latvians. *Biopolym Cell*. 2014; **30**(4):305–9.
46. *Jean-Philippe J, Paz S, Caputi M*. hnRNP A1: the Swiss army knife of gene expression. *Int J Mol Sci*. 2013;**14**(9):18999– 9024.
47. *Guo R, Li Y, Ning J, Sun D, Lin L, Liu X*. HnRNP A1/A2 and SF2/ASF regulate alternative splicing of interferon regulatory factor-3 and affect immunomodulatory functions in human non-small cell lung cancer cells. *PLoS One*. 2013;**8**(4):e62729.
48. *Wu ZH, Shi Y.* When ubiquitin meets NF-κB: a trove for anti-cancer drug development. *Curr Pharm Des*. 2013;**19**(18):3263–75.
49. *Dozmorov M, Wu W, Chakrabarty K, et al*. 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*. 2009**;9**:152.
50. *Ozaki K, Sato H, Iida A, et al.* A functional SNP in PSMA6 confers risk of myocardial infarction in the Japanese population. *Nat Genet*. 2006; (8):921–5.
51. *Sjalander A, Birgander R, Saha N, Beckman L, Beckman G*. p53 polymorphisms and haplotypes show distinct differences between major ethnic groups. *Hum Hered*. 1996;**46**(1):41–8.
52. *Kopp A.* Dmrt genes in the development and evolution of sexual dimorphism. *Trends Genet*. 2012;**28**(4):175–84.
53. *Imhof A, Wolffe AP.* Transcription: gene control by targeted histone acetylation. *Curr Biol.* 1998;**8**(12):R422–4.
54. *Johnson CA, Turner BM.* Histone deacetylases: complex transducers of nuclear signals. *Semin Cell Dev Biol.* 1999;**10**(2):179–88.
55. *Adcock IM, Ito K, Barnes PJ.* Histone deacetylation: an important mechanism in inflammatory lung diseases. *COPD*. 2005;**2**(4):445–55.
56. *Chiappara G, Chanez P, Bruno A, et al.* Variable p-CREB expression depicts different asthma phenotypes. *Allergy*. 2007;**62**(7):787–94.
57. *Viemari JC.* Noradrenergic modulation of the respiratory neural network. *Respir Physiol Neurobiol*. 2008;**164**(1–2):123–30.
58. *Clark RI, Tan SW, Pean CB, et al.* MEF2 is an *in vivo* immune-metabolic switch. *Cell*. 2013;**155**(2):435–47.

Received 01.09.14

PUBLICATION IV

3.4. Association between the *PSMB5* and *PSMC6* genetic variations and children obesity in the Latvian population

ISSN 0233–7657. Biopolymers and Cell. 2014. Vol. 30. N 6. P. 477–480 doi: http://dx.doi.org/10.7124/bc.0008C3

BIOMEDICINE

UDC 577.21 + 616.248

Association between the *PSMB5* and *PSMC6* genetic variations and children obesity in the Latvian population

N. Paramonova[1], S. Kupca[2, 3], I. Rumba-Rozenfelde[2, 3], N. Sjakste[2, 4], T. Sjakste[1]

[1]Genomics and Bioinformatics, Institute of Biology of the University of Latvia
3, Miera Str., Salaspils, Latvia, LV2169

[2]Faculty of Medicine, University of Latvia
19, Raina Bulv., Riga, Latvia, LV1586

[3]University Children Hospital
20, Juglas Str., Riga, Latvia, LV-1079

[4]Latvian Institute of Organic Synthesis
21, Aizkraukles Str., Riga, Latvia, LV1006

natasa@email.lubi.edu.lv

According to the recent data the ubiquitin-proteasome system (UPS) is implicated in the pathogenesis of obesity. Aim of our study was to evaluate a possible association between genetic variations in the PSMB5 and PSMC6 genes and childhood obesity in the Latvian population. ***Methods****. The rs11543947 (PSMB5), rs2295826 and rs2295827 (PSMC6) were genotyped in 94 overweight children versus 191 controls. Stratification was made by family history and sex.* ***Results****. Heterozygous genotype at rs11543947 (PSMB5) manifested association with the disease ($P < 0.01$) in total group and in patients with family history (OR = 2.445 [95 % CI 1.378–4.339] and OR = = 2.746 [95 % CI 1.427–5.283], respectively). This genotype was observed more frequently ($P < 0.05$) in males with family obesity and in females without family history ($P < 0.01$).The heterozygotes at rs2295826 and rs2295827 showed association ($P < 0.01$) in obesity (OB), in patients with family history (OR = 2.119 [95 % CI 1.207–3.718] and OR = 2.379 [95 % CI 1.249–4.533], respectively) and in males group. The rs11543947/ rs2295826-rs2295827 multi locus genotype heterozygous at all the studied loci and the haplotype represented by the rare alleles were more frequent in obese children when compared to controls ($P < 0.001$ and $P = 0.0001$ respectively).* ***Conclusions****. Genetic variations of the PSMB5 (rs11543947) and PSMC6 (rs2295826 and rs2295827) genes can influence childhood obesity in Latvians.*

Keywords: PSMB5, PSMC6, SNPs, obesity, familial obesity, genotype–sex interaction.

Introduction. The ubiquitin-proteasome system (UPS) has been recently shown to be implicated in the pathogenesis of obesity (OB). It has been demonstrated that plasma ubiquitin and proteasome levels inversely correlated with a male body mass index in Southern Taiwan and Japanese population [1, 2] and that the proteasome dysfunction mediated the obesity-induced endoplasmic reticulum stress and insulin resistance in the liver [3]. A significant association was found between genetic variants of the *PSMD9* gene and obesity-associated phenotypes in Italians [4]. The mutation in the *PSMB8* gene has been reported to be associated with the autoinflammatory syndrome with lipodystrophy in Japanese [5]. Earlier we have detected the association of the *PSMA3* gene polymorphisms with susceptibility to obesity in Latvian children [6].

The current study was aimed to elucidate whether the single nucleotide polymorphisms (SNPs) of the *PSMB5* (rs11543947) and *PSMC6* (rs2295826 and rs2295827) genes are associated with the children obesity in the Latvian population.

Materials and methods. Case/control groups (94 obese children and 191 controls, respectively) were described previously [6, 7]. The rs2295826 and rs2295827

genotyping data for controls were extracted from Sjakste *et al*. [7] and used in current study for the association analysis. The study was approved by the Central Medical Ethics Commission of the Latvian Ministry of Health.

The *PSMB5* (rs11543947) and *PSMC6* (rs2295826, rs2295827) genotyping procedures and analysis of amplified and digested products were the same as published previously [7, 8].

Single loci (SLGs) and multi locus rs11543947/ rs2295826-rs2295827 genotypes and alleles' frequencies were estimated by direct gene counting. The deviations from the Hardy-Weinberg equilibrium and association/correlation analyses were tested by the χ^2 test (allele, multi-loci genotype frequencies, recessive and dominant, over dominant and multiplicative models) using PAST (version 3.03) software for Windows [9]. Only $2 < OR < 0.5$ was considered to be clinically significant. DnaSP version 5 (http://www.ub.es/dnasp, [10]) was used to reconstruct the haplotypes from un-phased genotypes, and to evaluate the nucleotide and haplotype genetic diversity and pairwise linkage disequilibrium (LD) between the loci (D' and r^2). Stratification was performed by the sex and obesity family history.

Results and discussion. In both cases and controls genotyping call rate was 100 %. Homozygotes on common alleles for all loci were observed more frequently in controls than obese patients (more than 80 % and less than 70 %, respectively) and the tendencies were similar in children with and without family history (Table). The dominant and multiplicative models [11] were the most informative to reveal OB susceptible genotypes and alleles, respectively.

The rs11543947 heterozygous genotype showed a modest ($P < 0.01$) association with OB risk in total group and in children with familial OB. A rare T allele was observed more frequently ($P < 0.05$) in both these groups than in controls. The rs2295826 and rs2295827 loci alleles in OB patients were found to be in full linkage ($D' = 1$; $r^2 = 1$), as it was previously reported for Latvian controls [7]. These loci heterozygous genotypes were found to be in association ($P < 0.01$) with OB and familial OB groups. Rare alleles manifested nominal association with the disease ($P < 0.05$).

The family background was found to be an obesity risk factor (OR = 1.86 for mother; OR = 2.98 for siblings) among Chinese male youths [12]. A significant difference ($p < 0.01$) was observed between the subjects having the obesity family history and non-obese controls in rs2348071 heterozygous genotype frequencies at the *PSMA3* locus [6]. Bennet *et al*. have shown [13] that a greater predisposition to diabetes in Middle Eastern immigrants may be explained by a more extensive family history of the disorder.

In the current association study an interaction between sexes and familial obesity was revealed. The rs11543947 risk CT genotype and rare T allele were found more frequently in males with family obesity ($P < 0.05$) and in females without family history ($P < 0.01$). The rs2295826 and rs2295827 heterozygous genotypes and minor alleles were associated ($P < 0.01$ and $P < 0.05$, respectively) with all OB subgroups in males.

Sex specific differences in incidence and severity are also well known features for epidemiology of obesity. The sex-influenced association of obesity with genetic variations at the *LYPLAL1* locus, which encodes a lipase/esterase expressed in adipose tissue was suggested [14]. Seven new loci exhibited marked sexual dimorphism with a stronger effect on weight-hip ratio in women than men (P for sex difference < 0.05 to $P < 0.0001$) [15].

Thus, our data on the importance of interactions between the family history and sex for susceptibility for childhood obesity confirm the previous reports. Obviously having in our disposition a collection of a relatively small number of samples we can't reach high statistical reliability of results. The small number of subjects and the wide range of values (reflected by large standard deviation) precluded reliable statistical confirmation of this admission. However, even a study with small subject number reveals the trend on the association of the disease and can predict the common trends of association also for larger sample groups collected in larger populations.

The rs11543947/rs2295826-rs2295827 three locus genotype heterozygous at all the loci involved was found about four times more frequently in OB patients than in controls (about 3 and 13 %, respectively) and showed to be OB susceptible ($P < 0.001$). The multi locus T/G-T haplotype represented by rare alleles of the studied loci was more frequent ($P = 0.0001$) in obese children when compared to controls (0.5 and 6 %, respectively).

SNPs association with childhood obesity in common cohort, with and without family history groups

Gene/SNP ID	MA/Genotype	Number (frequency %)				Statistically significant association		
		Control, *n* =191	OB, *n* = 94	FH, *n* = 59	NFH, *n* = 32	*P*	OR	[95 % CI]
PSMB5/rs11543947	T	34 (8.90)	31 (16.49)	21 (17.80)	10 (15.63)	$< 0.05^{OB}$	2.021	[1.247–3.391]
		–	–	–	–	$< 0.05^{FH}$	2.216	[1.236–3.971]
		–	–	–	–	$> 0.05^{NFH}$	–	–
	CC	159 (83.25)	63 (67.02)	38 (64.41)	22 (68.75)	–	–	–
	CT	30 (15.71)	31 (32.98)	21 (35.59)	10 (31.25)	$< 0.01^{OB}$	2.445	[1.378–4.339]
		–	–	–	–	$< 0.01^{FH}$	2.746	[1.427–5.283]
	TT	2 (1.04)	No	No	No	–	–	–
PSMC6/rs2295826-rs2295827 (D' = 1; $r^2 = 1$)*	G	40 (10.47)	33 (17.55)	22 (18.64)	10 (15.63)	$< 0.05^{OB}$	1.820	[1.109–2.987]
		–	–	–	–	$< 0.05^{FH}$	1.959	[1.116–3.439]
		–	–	–	–	$> 0.05^{NFH}$	–	–
	AA	155 (81.15)	63 (67.02)	38 (64.41)	23 (71.88)	–	–	–
	AG	32 (16.75)	29 (30.85)	20 (33.90)	8 (25.00)	$< 0.01^{OB}$	2.119	[1.207–3.718]
		–	–	–	–	$< 0.01^{FH}$	2.379	[1.249–4.533]
		–	–	–	–	$> 0.05^{NFH}$	–	–
	GG	4 (2.10)	2 (2.13)	1 (1.69)	1 (3.13)	–	–	–
Loci	L1/L2-L3	–	–	–	–	–	–	–
Heterozygous genotypes	CT/AG-CT	5 (2.62)	12 (12.77)	7 (11.86)	5 (15.63)	$< 0.001^{OB}$	9.018	[2.967–27.410]
Rare alleles haplotype	T/G-T	2 (0.52)	11 (5.85)	7 (5.93)	4 (6.25)	0.0001^{OB}	13.660	[2.895–62.524]

N o t e. OB patients were represented by 59 children with obesity family history, 32 children without family history and there were no data on familial history of obesity in 3 cases. Abbreviations: OB – Obesity; FH – obesity with family history; NFH – obesity without family history. *Alleles and genotypes frequencies of the rs2295826 and rs2295827 were found in full linkage. L1/L2-L3 corresponds to the rs11543947 and rs2295826-rs2295827 loci respectively.

Susceptibility of genes encoding proteasome subunits to the immunity related disorders was studied previously in Latvians. The rs1048990 (*PSMA6*), rs2295826 and rs2295827 (*PSMC6*), rs2348071 (*PSMA3*) were found to be associated with susceptibility to bronchial asthma [16]. The rs2277460 (*PSMA6*), rs2295826 and rs2295827 (*PSMC6*) and rs2348071 (*PSMA3*) manifested association with juvenile idiopathic arthritis [8]. The rs2348071 (*PSMA3*) was found to be associated with multiple sclerosis [17] and childhood obesity [6].

Obesity in the young characterized by visceral fat accumulation has been shown to be a major risk factor for adult-onset type 2 diabetes mellitus [18, 19], therefore, we cannot exclude in our cases a possibility of the development of diabetes in adult age.

Conclusions. The genetic variation of *PSMB5* (rs11543947) and *PSMC6* (rs2295826 and rs2295827) 14q proteasomes subunits can influence OB in Latvians.

Acknowledgements & Funding. This study was supported by the Latvian National Research Program «Biomedicine 2014».

Асоціація між генетичними варіантами *PSMB5* і *PSMC6* та дитячим ожирінням з-поміж жителів Латвії

Н. Парамонова, С. Купча, И. Румба-Розенфелде, Н. Сьяксте, Т. Сьяксте

Резюме

Згідно з останніми даними, убіквітин-залежна протеасомна система бере участь у патогенезі ожиріння. ***Мета****. Оцінити можливий зв'язок між генетичними варіантами протеасомних генів PSMB5 і PSMC6 та схильністю до захворювання дитячим ожирін-*

*ням у Латвійській популяции. **Методи**. Локуси rs11543947 (PSMB5), rs2295826 і rs2295827 (PSMC6) генотипували у 94 дітей з надлишковою вагою і у 191 здорового індивіда. Оцінку проводили за асоціацією з ожирінням як таким, за сімейною історією та за статтю. **Результати**. Гетерозиготний генотип, який належить до локусу rs11543947 (PSMB5), виявився помірно асоційованим (P < 0,01) із захворюванням як таким і з ожирінням з сімейною історією (співвідношення шансів СШ = 2,445 [95 % ДІ 1.378–4.339] і СШ = 2,746 [95 % ДІ 1.427–5.283] відповідно). Цей генотип найчастіше спостерігався у чоловіків із сімейною історією ожиріння (P < 0,05) та у жінок без сімейної історії (P < 0,01). Гетерозиготні генотипи по локусах rs2295826 і rs2295827 знайдено в помірній асоціації (P < 0,01) в основній групі захворювання і у пацієнтів з сімейною історією (СШ = 2,119 [95 % ДІ 1.207–3.718] і СШ = = 2,379 [95 % ДІ 1.249–4.533] відповідно), а також у чоловіків. Багатолокусний генотип rs11543947/rs2295826-rs2295827, представлений гетерозиготами по всіх локусах, і гаплотип, представлений рідкісними алелями, були найчастішими у групі хворих на ожиріння порівняно з контрольною групою (P < 0,001 і P = 0,0001 відповідно). **Висновки**. Генетичні варіації локусів PSMB5 (rs11543947) і PSMC6 (rs2295826 і rs2295827) можуть впливати на схильність до захворювання ожирінням у детей Латвійської популяції.*

Ключові слова: PSMB5, PSMC6, SNPs, ожиріння, сімейне ожиріння, взаємодія генотип–стать.

Ассоциация между генетическими вариантами *PSMB5* и *PSMC6* и детским ожирением среди жителей Латвии

Н. Парамонова, С. Купча, И. Румба-Розенфелде,
Н. Сьяксте, Т. Сьяксте

Резюме

*Согласно последним данным, убиквитин-зависимая протеасомная система участвует в патогенезе ожирения. **Цель**. Оценить возможную связь между генетическими вариантами протеасомных генов PSMB5 и PSMC6 и подверженностью заболеванию детским ожирением в Латвийской популяции. **Методы**. Локусы rs11543947 (PSMB5), rs2295826 и rs2295827 (PSMC6) генотипировали у 94 детей с избыточным весом и у 191 здорового индивида. Оценку проводили по ассоциации с ожирением как таковым, с семейной историей и с полом. **Результаты**. Гетерозиготный генотип, относящийся к локусу rs11543947 (PSMB5), оказался умеренно ассоциированным (P < 0,01) с заболеванием как таковым и с ожирением с семейной историей (отношение шансов ОШ = = 2,445 [95 % ДИ 1.378–4.339] и ОШ = 2,746 [95 % ДИ 1.427–5.283] соответственно). Этот генотип наиболее часто наблюдался у мужчин с семейной историей ожирения (p < 0,05) и у женщин без семейной истории (P < 0,01). Гетерозиготные генотипы по локусам rs2295826 и rs2295827 найдены в умеренной ассоциации (P < 0,01) в основной группе заболевания и у пациентов с семейной историей (ОШ = 2,119 [95 % ДИ 1.207–3.718] и ОШ = = 2,379 [95 % ДИ 1.249–4.533] соответственно) и у мужчин. Многолокусный генотип rs11543947/rs2295826-rs2295827, представленный гетерозиготами по всем локусам, и гаплотип, представленный редкими аллелями, были наиболее частыми в группе больных ожирением по сравнению с контрольной группой (P < 0,001 и P = 0,0001 соответственно). **Выводы**. Генетические вариации локусов PSMB5 (rs11543947) и PSMC6 (rs2295826 и rs2295827) могут влиять на подверженность заболеваемости ожирением у детей в Латвийской популяции.*

Ключевые слова: PSMB5, PSMC6, SNPs, ожирение, семейное ожирение, взаимодействие генотип–пол.

REFERENCES

1. *Chang TL, Chang CJ, Lee WY, Lin MN, Huang YW, Fan K*. The roles of ubiquitin and 26S proteasome in human obesity. *Metabolism*. 2009;**58**(11):1643–8.
2. *Sakamoto K, Sato Y, Sei M, Ewis AA, Nakahori Y*. Proteasome activity correlates with male BMI and contributes to the differentiation of adipocyte in hADSC. *Endocrine*. 2010;**37**(2):274–9.
3. *Otoda T, Takamura T, Misu H, et al*. Proteasome dysfunction mediates obesity-induced endoplasmic reticulum stress and insulin resistance in the liver. *Diabetes*. 2013;**62**(3):811–24.
4. *Gragnoli C*. Overweight condition and waist circumference and a candidate gene within the 12q24 locus. *Cardiovasc Diabetol.* 2013;**12**:2.
5. *Kitamura A, Maekawa Y, Uehara H, et al*. A mutation in the immunoproteasome subunit PSMB8 causes autoinflammation and lipodystrophy in humans. *J Clin Invest.* 2011;**121**(10):4150–60.
6. *Kupca S, Sjakste T, Paramonova N, et al*. Association of obesity with proteasomal gene polymorphisms in children. *J Obes*. 2013; **2013**:638154.
7. *Sjakste T, Paramonova N, Wu LS-S, et al.* PSMA6 (rs2277460, rs1048990), PSMC6 (rs2295826, rs2295827) and PSMA3 (rs2348071) genetic diversity in Latvians, Lithuanians and Taiwanese. *Meta Gene*. 2014;**2**:283–98.
8. *Sjakste T, Paramonova N, Rumba-Rozenfelde I, Trapina I, Sugoka O, Sjakste N*. Juvenile idiopathic arthritis subtype- and sex-specific associations with genetic variants in the PSMA6/PSMC6/PSMA3 gene cluster. *Pediatr Neonatol*. 2014;**55**(5):393–403.
9. *Hammer Q, Harper D, Ryan P*. Past: Paleontological Statistics software packarge for education and data analysis. *Paleontologija Electronica*. 2001; **4**(1):9pp
10. *Rozas J*. DNA sequence polymorphism analysis using DnaSP. *Methods Mol Biol*. 2009;**537**:337–50.
11. *Lewis CM*. Genetic association studies: design, analysis and interpretation. *Brief Bioinform*. 2002;**3**(2):146–53.
12. *Shi H, Jiang B, Wei Sim JD, Chum ZZ, Ali NB, Toh MH*. Factors associated with obesity: a case-control study of young adult Singaporean males. *Mil Med.* 2014;**179**(10):1158–65.
13. *Bennet L, Lindblad U, Franks PW*. A family history of diabetes determines poorer glycaemic control and younger age of diabetes onset in immigrants from the Middle East compared with native Swedes. *Diabetes Metab.* 2014. pii: S1262-3636(14)00136-0.
14. *Benjamin AM, Suchindran S, Pearce K, et al.* Gene by sex interaction for measures of obesity in the framingham heart study. *J Obes*. 2011;**2011**:329038.
15. *Heid IM, Jackson AU, Randall JC, et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet*. 2010;**42**(11):949–60.
16. *Paramonova N, Wu L. Shi-Shin, Rumba-Rozenfelde I, Wang J-Y, Sjakste N, Sjakste T*. Genetic variants in the PSMA6, PSMC6 and PSMA3 genes associated with childhood asthma in Latvian and Taiwanese populations. *Biopolym Cell*. 2014; **30**(5):377–87.
17. *Kalnina J, Paramonova N, Sjakste N, Sjakste T*. Study of association between polymorphisms in the PSMB5 (rs11543947) and PSMA3 (rs2348071) genes and multiple sclerosis in Latvians. *Biopolym Cell*. 2014; **30**(4):305–9.
18. *Kaufman F*. Type 2 diabetes mellitus in children and adolescents: the new epidemic. *Obesity in Childhood and Adolescence*. Eds W. Keiss, C. Marcus, M. Wabitsch. New Jersey, Karger publ., 2004: 183–4.
19. *de Luca C, Olefsky JM*. Stressed out about obesity and insulin resistance. *Nat Med.* 2006;**12**(1):41–2.

Received 01.10.14

PUBLICATION V

3.5. Functional significance of microsatellite markers

REVIEW

Medicina (Kaunas) 2013;49(12):505-9

Functional Significance of Microsatellite Markers

Tatjana Sjakste[1], Natalia Paramonova[1], Nikolajs Sjakste[2]
[1]Genomics and Bioinformatics, Institute of Biology, University of Latvia,
[2]Faculty of Medicine, University of Latvia, Latvia

Key Words: *microsatellite; human diseases; promoter; intron; exon.*

Summary. *The review summarizes literature data on the positive results of association studies between the length of microsatellite repeats and predisposition to pathologies. Actually, the data can be classified according to the localization of the microsatellite: in the gene promoter, in the part of exon 1 coding the signal sequence, in gene introns, in the coding areas of genes, and in 3'-untranslated regions. The functional significance of microsatellite length changes can be evaluated in many cases. The authors came up to the conclusion that further studies on microsatellite associations with diseases remain prospective as they reflect changes in the gene functional activity.*

Introduction

Microsatellites (MSs) are repeating sequences of 2–6 base pairs of DNA (1). They are used as molecular markers in genetics for kinship, population, and other studies. They can also be used for the studies of gene duplication or deletion, marker-assisted selection, and fingerprinting. Microsatellites are distributed throughout the genome (1). Being variable genetic elements, microsatellites provide a potent tool for the individual characterization of genomes. Variability is generated due to replication slippage caused by mismatches between DNA strands while being replicated during meiosis (2), and the event can occur once per 1000 generations (3). This slippage is much more common compared with point mutations (4). Microsatellite repeats mutagenize human genomes and alter the human genomic landscape across generations (5). The utility of microsatellites has been demonstrated by the study comprising 2058 germline changes discovered by analyzing 85 289 Icelanders at 2477 microsatellites. The paternal-to-maternal mutation rate ratio is 3.3, and the rate in fathers doubles from the age 20 to 58, whereas there is no association with age in mothers. Longer microsatellite alleles are more mutagenic and tend to decrease in length, whereas the opposite is seen for shorter alleles (6).

Microsatellites remain highly informative and useful measures of genomic variation for linkage and association studies despite the fact that general preference is given to single-nucleotide polymorphisms (SNPs). Microsatellites are much more genetically diverse compared to SNPs; they generate a greater haplotype diversity (7).

Although mostly used as structural genetic markers, microsatellites perform several functions in the genome, which are still far from being completely understood. However, actually it has become clear that MSs and their flanking regions are involved in multiple gene and genome functions. MSs are known to form nuclear matrix anchorage sites (8), tissue-specific matrix attachment sites (9), and binding sites with vimentin and glial fibrillary acidic protein (10) and give rise to complex DNA spatial structures of extreme functional significance (10–12). MSs appear to be important components of insulators (13), silencers (12), and enhancers (14). MSs are also involved in the regulation of alternative splicing (15), mRNA stability (16), and recombination and repair (17, 18).

Expansions of microsatellite DNA repeats cause nearly 30 developmental and neurological inherited disorders (19). Further, we shall try to summarize and classify the data about the association between MSs and human diseases in connection to the localization of repeats in the genes and their possible functional role. Triplet expansion diseases will be excluded from our analysis, as the problem has been extensively reviewed (20, 21).

Promoter Microsatellites. MSs can determine the activity of the upstream gene regulation elements like the locus control region of the beta-globin gene domain. The (AT)(8)N(12)GT(AT)(7) configuration of a microsatellite found in the hypersensitive site of the structure is associated with a special form of sickle cells, Tunisian βs chromosomes (22). The most common allele of the MS marker in STAT4, the STAT4-MS1-254 allele, located in the 5' flanking region of the gene, is significantly associated with sarcoidosis (23). Changes in the length of microsatellites within promoters and other cis-regulatory regions can also change the level of gene expression, and they are linked to abundant variations in cis-regulatory control regions in the human genome (24). For example, a CA-repeat mi-

Correspondence to N. Sjakste, Faculty of Medicine, University of Latvia, Raiņa bulvāris 19, 1586 Riga, Latvia
E-mail: nikolajs.sjakste@lu.lv

crosatellite in the insulin-like growth factor 1 promoter is associated with the level of this growth factor. It has turned out that the intensity of the gene transcription is regulated by the interaction of several SNPs and microsatellite-generating haplotypes with lower or higher levels of the gene transcription (25). Promoter microsatellites tend to be guanine-cytosine rich; they are often found at the start of genes and are probably associated with the regulatory elements such as CpG islands, G-quadruplexes (G4), and untranslated regulatory regions. Numerous promoter microsatellites possess the potential to influence human phenotypes by generating mutations in regulatory elements, which may ultimately lead to a disease (26). A CpG-CA repeat within the human endothelin-converting enzyme 1 promoter is highly polymorphic, harbors transcriptional start sites, is able to recruit the transcription factors and poly(ADP-ribose) polymerase-1 and splicing factors, and is functional regarding haplotype-specific promoter activity. The overall CpG-CA repeat composition of patients with Alzheimer's disease and nondemented control individuals has been found to be distinct (27). A length polymorphism of GT repeats in the promoter region of the human heme oxygenase-1 (*HO-1*) gene modulates the transcription of this gene (28). Numerous studies have linked human *HO-1* gene promoter polymorphisms to a risk of vascular diseases (29). Persons carrying longer (GT)(n) repeats in the *HMOX1* gene (L allele) promoter may be at a higher risk of type 2 diabetes mellitus (30). Functional analyses have shown that the persons with impaired glucose regulation and type 2 diabetes mellitus, carrying the L/L (GT)(n) genotype, have significantly lower HO-1 protein expression levels than those with the S/S genotype (31). The same microsatellite is associated with susceptibility to cardiovascular complications of the disease. Patients with longer lengths of GT repeats in the heme oxygenase-1 gene promoter exhibit higher inflammation and oxidative stress. These patients have a higher risk of long-term cardiovascular events and mortality (32). A short allele of the same microsatellite might be associated with an abdominal aortic aneurysm (33). Long (GT)n repeats in the microsatellite polymorphism region of the *HMOX1* gene have been reported to be associated with symptomatic malaria (34).

The aldose reductase (*AKR1B1*) gene promoter harbors a (CA)n microsatellite significantly associated with diabetic retinopathy. The z-2 microsatellite has been found to confer risk in type 1 and type 2 diabetes mellitus and the z+2 microsatellite to confer protection against diabetic retinopathy in type 2 diabetes mellitus regardless of ethnicity (35, 36). The S allele of a ((CCTTT)(n) repeat in the promoter of the NOS2 gene is associated with both hypertension and responsiveness to antihypertensive drug therapy (37). The same microsatellite is also associated with diabetic retinopathy, as well as the (GT)n promoter repeat in the tumor necrosis factor β gene (36).

Promoter MSs might also be associated with mental problems; the promoter TA microsatellite repeat in the estrogen receptor alpha gene is significantly associated with postpartum depression (38). The arginine vasopressin receptor 1A gene (*AVPR1A*) is widely expressed in the brain and is considered to be a key receptor in the regulation of social behavior. 5'-Flanking region polymorphisms in the human AVPR1A, RS3, and RS1 show differences in relative promoter activity by length. Shorter repeat alleles of RS1 and RS3 have decreased relative promoter activity in the human neuroblastoma cell line SH-SY5Y. The short alleles of RS1 are associated with autism (39).

A recent meta-analysis by Shen et al. has reported that CYP11A1 promoter microsatellite [TTTA] n repeat polymorphisms may contribute to increasing susceptibility to the risk of polycystic ovary syndrome (40).

Signal Sequence Microsatellites. Some microsatellites are localized in translated areas of the genes. For example, the carnosinase gene contains a D18S880 microsatellite formed of a leucine triplet repeat in its signal sequence. Homozygotes for 5 trinucleotide repeats in this microsatellite are susceptible to diabetic nephropathy (41). The human signal transducer and activator of transcription 6 (*STAT6*) gene represents one of the most promising candidate genes for asthma and other inflammatory diseases on the chromosomal region 12q13-q24. The gene exon 1 contains a GT repeat upstream the first methionine codon. Allele A4 of the GT repeat polymorphism is associated with an increase in the eosinophil cell count (42). The genotype of 13/15-GT repeat allele heterozygosity is significantly associated with allergic subjects (43).

Microsatellites of Coding Regions. Besides trinucleotide expansion diseases, characterized mostly by polyglutamine tracts (poly-Q), which cannot be analyzed here due to space limitations, an interesting trinucleotide repeat has been identified in the MIC-A gene. The exon 5 microsatellite polymorphism of the MIC-A gene consists of 5 alleles based on the number of GCT triplet repeat units (alleles A4, A5, A6, and A9) and the presence of an additional nucleotide insertion (allele A5.1). CGT repeats regulate the number of Ala residues in the protein, and the A5.1 leads to a frameshift mutation. The exon encodes the membrane-binding domain of the protein (44). The microsatellite alleles are associated with Addison's disease (44) and type 1 diabetes mellitus (45–47). Some alleles are protective against

juvenile idiopatic arthritis (48). Variations of CAG (Gln) repeats in the androgen receptor gene in physiological limits, not causing insensitivity to androgens, can influence certain physiological parameters. Shorter androgen receptor (AR) CAG is associated with low HDL-C and testosterone levels (49).

Intronic Microsatellites. Microsatellites within introns also influence a phenotype, through ways that are not currently understood; this is the cause of numerous associations of microsatellite repeat polymorphisms with human diseases. For example, a GAA triplet expansion in the first intron of the X25 gene appears to interfere with transcription and causes Friedreich ataxia (50). Subjects having more CA repeats in the first intron of the type 2 11β-hydroxysteroid dehydrogenase gene (HSD11B2) are susceptible to developing abnormal glucose tolerance (51). A repeat polymorphism in the fourth intron of the NOS3 gene is linked to hypertension (52). We have detected an association between type 2 diabetes mellitus and microsatellite markers of the region 14q13 localized in the introns of the PSMA6 and KIAA0391 genes, rs63749745, rs71444202, and rs34580276(53).

Three microsatellite loci, i.e., (ATCC)n1, D1S1621, and (ATCC)n2, in the DISC1 gene show a significant association with schizophrenia. The microsatellites occur in intronic sequences in the vicinity of a critical splice junction that gives rise to the expression of the DISC1 isoforms (54).

Intronic microsatellite polymorphisms determine susceptibility to certain neoplasias. For example, polymorphisms in the CT dinucleotide repeat in intron 3 of the transcription factor GATA3 gene are associated to a certain extent with the risk of breast cancer, i.e., women who carry 17-CT or 18-CT alleles of the *GATA3* gene are at a lower risk of developing breast cancer (55). The polymorphic dinucleotide CA tandem repeat (ESR2_CA), located in intron 5 of the estrogen receptor gene 2 gene ESR2 (14q23.2), is associated with the risk of breast cancer in African women (56). Intronic D19S884 marker A7 allele of the fibrillin 3 gene is associated with polycystic ovary syndrome (57).

Intronic microsatellites repeats are implicated in the pathogenic mechanisms of several autoimmune diseases. The *SLC26A4* gene, involved in the genetic susceptibility of autoimmune thyroid disease, harbors 2 microsatellites in introns 10 and 20, and longer alleles of these markers appear to be associated with Hashimoto thyroiditis (58). The intronic rs63749745 marker of the 14q13 locus has manifested a high level of association with Graves' disease (59), and rs71928782, rs5807818, rs71444202, and rs345802276 have been found to be in association with juvenile idiopathic arthritis in children (60).

Polymorphisms present in the first intron of IFN-γ may have an important role in the regulation of the immune response, which could have functional consequences for gene transcription. The microsatellite encoding 16 CA repeats has been shown to be significantly associated with the paucibacillary form of lepra compared with multibacillary patients (61). The microsatellite marker IFNGR2-MS1, located in the 50-upstream region of the interferon gamma receptor 2 gene (*IFNGR2*), shows a significant association with tuberculosis (62).

One allele of the D6S1276 microsatellite in intron 1 of the BMP5 gene is associated with the risk of osteoarthritis in women; 2 alleles are protective (63).

3'-UTR Microsatellites. Microsatellites localized in the 3'-untranslated regions (3'-UTRs) may affect the final mRNA stability, the localization, the export from the nucleus and the translation efficiency. The androgen receptor CAG repeat polymorphism (AR CAG) affects receptor transcriptional activity (the shorter repeats, the more sensitive AR) and is associated with androgenic parameters and obesity (49). The conserved regulatory sequences within the 3'-UTRs and the specific elements binding to them enable gene expression control at the posttranscriptional level, and all these processes reflect the actual state of the cell (64). Shorter alleles of the microsatellites in the 3' flanking region of the leptin gene, coding for a protein hormone, mainly synthesized in adipocytes, which regulates the food intake and energy expenditure of the body, are significantly associated with hypertension (65). Reduced repeat lengths in the EGFR gene 3'-UTR polyA repeat are linked with osteosarcomas (66). Several alleles of the microsatellite (AT)n in the 3'-UTR of anticytotoxic T lymphocyte antigen-4 (CTLA-4) gene, namely 104-, 106-, 110-, and 116-bp alleles, were observed to be predisposing to recurrent miscarriage (67).

Remote and Locus-Specific Microsatellites. In some cases, the association with diseases is found for microsatellites localized far from the candidate genes. Marker D12S96 is localized 5.653 cM downstream the vitamin D receptor (VDR) gene. Despite this long distance and obscure functional relations, statistically significant linkage disequilibrium has been detected between allele 22 of locus D12S96 and osteoporosis (68). In some cases, the association between microsatellites and candidate genes is not traced at all as these are sooner the locus than gene markers. The 8p21-23 region microsatellites D8S136 and D8S520 are consistently and strongly related with prostate cancer (69). The locus has been traced due to a frequent loss of heterozigocity in tumors, but not as a result of association studies. The D1S2726 microsatellite, located 30 kb from the *KCNA3* gene, which encodes the voltage-gated potassium channel Kv1.3, is associated with susceptibility to autoimmune pancreatitis (70).

Conclusions

The above data clearly indicate that most microsatellites manifesting the association with human pathologies are harbored in genes encoding enzymes involved in pathogenesis of the pathologies; in many cases, the impact of the changes in the microsatellite length on the gene function can be evaluated. Thus, further studies on microsatellite associations with diseases remain prospective, despite numerous whole-genome association studies. Contribution of studies with individual polymorphisms to the understanding of genetic background of diseases should not be underestimated.

Acknowledgments

The costs of this work were covered from the European Regional Development Foundation project No. 2010/0315/2DP/2.1.1.1.0/10/APIA/VIAA/026 and the scientific co-operation project No. 10. 0010 of the Latvian Council of Sciences "Genetic Study of Susceptibility to Diseases and Ageing in the Latvian Population" subproject "Wide Scanning of the Proteasomal Gene Polymorphism in the Latvian Population and Its Association With Autoimmune Diseases."

Statement of Conflict of Interest

The authors state no conflicts of interest.

References

1. King DG. Evolution of simple sequence repeats as mutable sites. Adv Exp Med Biol 2012;769:10-25.
2. Tautz D, Schlötterer C. Simple sequences. Curr Opin Genet Dev 1994;4:832-7.
3. Weber JL, Wong C. Mutation of human short tandem repeats. Hum Mol Genet 1993;2:1123-8.
4. Jarne P, Lagoda PJ. Microsatellites, from molecules to populations and back. Trends Ecol Evol 1996;11:424-9.
5. Grandi FC, An W. Non-LTR retrotransposons and microsatellites: partners in genomic variation. Mob Genet Elements 2013;3:e25674.
6. Sun JX, Helgason A, Masson G, Ebenesersdóttir SS, Li H, Mallick S, et al. A direct characterization of human mutation based on microsatellites. Nat Genet 2012;44:1161-5.
7. Gulcher J. Microsatellite markers for linkage and association studies. Cold Spring Harb Protoc 2012;2012:425-32.
8. Boulikas T. Nature of DNA sequences at the attachment regions of genes to the nuclear matrix. J Cell Biochem 1993; 52:14-22.
9. Lenartowski R, Goc A. Tissue-specific association of the human tyrosine hydroxylase gene with the nuclear matrix. Neurosci Lettr 2002;330:151-4.
10. Tolstonog GV, Li G, Shoeman RL, Traub P. Interaction in vitro of type III intermediate filament proteins with higher order structures of single-stranded DNA, particularly with G-quadruplex DNA. DNA Cell Biol 2005;24:85-110.
11. Li G, Tolstonog GV, Traub P. Interaction in vitro of type III intermediate filament proteins with Z-DNA and B-Z-DNA junctions. DNA Cell Biol 2003;22:141-69.
12. Rothenburg S, Koch-Nolte F, Haag F. DNA methylation and Z-DNA formation as mediators of quantitative differences in the expression of alleles. Immunol Rev 2001;184:286-98.
13. Filippova GN, Thienes CP, Penn BH, Cho DH, Hu YJ, Moore JM, et al. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat Genet 2001;28:335-43.
14. Bassuny WM, Ihara K, Sasaki Y, Kuromaru R, Kohno H, Matsuura, N, et al. A functional polymorphism in the promoter/enhancer region of the FOXP3/Scurfin gene associated with type 1 diabetes. Immunogenetics 2003;55:149-56.
15. Hui J, Hung LH, Heiner M, Schreiner S, Neumüller N, Reither G, et al. Intronic CA-repeat and CA-rich elements: a new class of regulators of mammalian alternative splicing. EMBO J 2005;24:1988-98.
16. Lee JH, Jeon MH, Seo YJ, Lee YJ, Ko JH, Tsujimoto Y, et al. CA repeats in the 3'-untranslated region of bcl-2 mRNA mediate constitutive decay of bcl-2 mRNA. J Biol Chem 2004;279:42758-64.
17. Wang G, Vasquez KM. Z-DNA, an active element in the genome. Front Biosci 2007;12:4424-38.
18. Chin JY, Schleifman EB, Glazer PM. Repair and recombination induced by triple helix DNA. Front Biosci 2007;12:4288-97.
19. Kim JC, Mirkin SM. The balancing act of DNA repeat expansions. Curr Opin Genet Dev 2013;23:280-8.
20. Wang YH. Chromatin structure of repeating CTG/CAG and CGG/CCG sequences in human disease. Front Biosci 2007;12:4731-41.
21. Fondon JW, Hammock EA, Hannan AJ, King DG. Simple sequence repeats: genetic modulators of brain function and behavior. Trends Neurosci 2008;31:328-34.
22. Ben Mustapha M, Moumni I, Zorai A, Douzi K, Ghanem A, Abbes S. Microsatellite and single nucleotide polymorphisms in the β-globin locus control region-hypersensitive Site 2: SPECIFICITY of Tunisian βs chromosomes. Hemoglobin 2012;36:533-44.
23. Tanaka G, Matsushita I, Ohashi J, Tsuchiya N, Ikushima S, Oritsu M, et al. Evaluation of microsatellite markers in association studies: a search for an immune-related susceptibility gene in sarcoidosis. Immunogenetics 2005;56:861-70.
24. Rockman MV, Wray GA. Abundant raw material for cis-regulatory evolution in humans. Mol Biol Evol 2002;19: 1991-2004.
25. Chen HY, Huang W, Leung VH, Fung SL, Ma SL, Jiang H, et al. Functional interaction between SNPs and microsatellite in the transcriptional regulation of insulin-like growth factor 1. Hum Mutat 2013;34:1289-97.
26. Sawaya S, Bagshaw A, Buschiazzo E, Kumar P, Chowdhury S, Black MA, et al. Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. PLoS One 2013;8:e54710.
27. Li Y, Seidel K, Marschall P, Klein M, Hope A, Schacherl J, et al. A polymorphic microsatellite repeat within the ECE-1c promoter is involved in transcriptional start site determination, human evolution, and Alzheimer's disease. J Neurosci 2012;32:16807-20.
28. Hu YF, Lee KT, Wang HH, Ueng KC, Yeh HI, Chao TF, et al. The association between heme oxygenase-1 gene promoter polymorphism and the outcomes of catheter ablation of atrial fibrillation. PLoS One 2013;8:e56440.
29. Wu ML, Ho YC, Yet SF. A central role of heme oxygenase-1 in cardiovascular protection. Antioxid Redox Signal 2011;15:1835-46.
30. Bao W, Song F, Li X, Rong S, Yang W, Wang D, et al. Association between heme oxygenase-1 gene promoter polymorphisms and type 2 diabetes mellitus: a HuGE review and meta-analysis. Am J Epidemiol 2010;172:631-6.
31. Song F, Li X, Zhang M, Yao P, Yang N, Sun X, et al. Association between heme oxygenase-1 gene promoter polymorphisms and type 2 diabetes in a Chinese population. Am J Epidemiol 2009;170:747-56.
32. Chen YH, Hung SC, Tarng DC. Length polymorphism in heme oxygenase-1 and cardiovascular events and mortality in hemodialysis patients. Clin J Am Soc Nephrol 2013;8:1756-63.
33. Gregorek AC, Gornik KC, Polancec DS, Dabelic S. GT microsatellite repeats in the heme oxygenase-1 gene promoter associated with abdominal aortic aneurysm in Croatian patients. Biochem Genet 2013;51:482-92.
34. Mendonça VR, Luz NF, Santos NJ, Borges VM, Gonçalves MS, Andrade BB, et al. Association between the haptoglobin and heme oxygenase 1 genetic profiles and soluble CD163 in susceptibility to and severity of human malaria. Infect Immun 2012;80:1445-54.

35. Abhary S, Hewitt AW, Burdon KP, Craig JE. A systematic meta-analysis of genetic association studies for diabetic retinopathy. Diabetes 2009;58:2137-47.
36. Uthra S, Raman R, Mukesh BN, Rajkumar SA, Kumari P, Lakshmipathy P, et al. Diabetic retinopathy: validation study of ALR2, RAGE, iNOS and TNFB gene variants in a south Indian cohort. Ophthalmic Genet 2010;31:244-51.
37. Oliveira-Paula GH, Lacchini R, Coeli-Lacchini FB, Junior HM, Tanus-Santos JE. Inducible nitric oxide synthase haplotype associated with hypertension and responsiveness to antihypertensive drug therapy. Gene 2013;515:391-5.
38. Pinsonneault JK, Sullivan D, Sadee W, Soares CN, Hampson E, Steiner M. Association study of the estrogen receptor gene ESR1 with postpartum depression – a pilot study. Arch Womens Ment Health 2013;16:499-509.
39. Tansey KE, Hill MJ, Cochrane LE, Gill M, Anney RJ, Gallagher L. Functionality of promoter microsatellites of arginine vasopressin receptor 1A (AVPR1A): implications for autism. Mol Autism 2011;2:3.
40. Shen W, Li T, Hu Y, Liu H, Song M. Common polymorphisms in the CYP1A1 and CYP11A1 genes and polycystic ovary syndrome risk: a meta-analysis and meta-regression. Arch Gynecol Obstet 2014;289:107-18.
41. Zhu JM, Wang B, Li J, Chen GM, Fan YG, Feng CC, et al. D18S880 microsatellite polymorphism of carnosinase gene and diabetic nephropathy: a meta-analysis. Genet Test Mol Biomarkers 2013;17:289-94.
42. Duetsch G, Illig T, Loesgen S, Rohde K, Klopp N, Herbon N, et al. STAT6 as an asthma candidate gene: polymorphism-screening, association and haplotype analysis in a Caucasian sib-pair study. Hum Mol Genet 2002;11:613-21.
43. Tamura K, Suzuki M, Arakawa H, Tokuyama K, Morikawa A. Linkage and association studies of STAT6 gene polymorphisms and allergic diseases. Int Arch Allergy Immunol 2003;131:33-8.
44. Gambelunghe G, Falorni A, Ghaderi M, Laureti S, Tortoioli C, Santeusanio F, et al. Microsatellite polymorphism of the MHC class I chain-related (MIC-A and MIC-B) genes marks the risk for autoimmune Addison's disease. J Clin Endocrinol Metab 1999;84:3701-7.
45. Kumar N, Sharma G, Kaur G, Tandon N, Bhatnagar S, Mehra N. Major histocompatibility complex class I chain related gene-A microsatellite polymorphism shows secondary association with type 1 diabetes and celiac disease in North Indians. Tissue Antigens 2012;80:356-62.
46. Novota P, Kolostova K, Pinterova D, Novak J, Weber P, Treslova L, et al. Association of MHC class I chain related gene-A microsatellite polymorphism with the susceptibility to T1DM and LADA in Czech adult patients. Int J Immunogenet 2005;32:273-5.
47. Bilbao JR, Martín-Pagola A, Calvo B, Perez de Nanclares G, Gepv-N, Castaño L. Contribution of MIC-A polymorphism to type 1 diabetes mellitus in Basques. Ann N Y Acad Sci 2002;958:321-4.
48. Nikitina Zake L, Cimdina I, Rumba I, Dabadghao P, Sanjeevi CB. Major histocompatibility complex class I chain related (MIC) A gene, TNFa microsatellite alleles and TNFB alleles in juvenile idiopathic arthritis patients from Latvia. Hum Immunol 2002;63:418-23.
49. Stanworth RD, Kapoor D, Channer KS, Jones TH. Dyslipidaemia is associated with testosterone, oestradiol and androgen receptor CAG repeat polymorphism in men with type 2 diabetes. Clin Endocrinol (Oxf) 2011;74:624-30.
50. Bidichandani SI, Ashizawa T, Patel PI. The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. Am J Hum Genet 1998;62:111-21.
51. Mune T, Suwa T, Morita H, Isomura Y, Takada N, Yamamoto Y, et al. Longer HSD11B2 CA-repeat in impaired glucose tolerance and type 2 diabetes. Endocr J 2013;60:671-8.
52. Jemaa R, Ben Ali S, Kallel A, Feki M, Elasmi M, Taieb SH, et al. Association of a 27-bp repeat polymorphism in intron 4 of endothelial constitutive nitric oxide synthase gene with hypertension in a Tunisian population. Clin Biochem 2009;42:852-6.
53. Sjakste T, Kalis M, Poudziunas I, Pirags V, Lazdins M, Groop L, et al. Association of microsatellite polymorphisms of the human 14q13.2 region with type 2 diabetes mellitus in Latvian and Finnish populations. Ann Hum Genet 2007;71:772-6.
54. Cao F, Zhang H, Feng J, Gao C, Li S. Association study of three microsatellite polymorphisms located in introns 1, 8, and 9 of DISC1 with schizophrenia in the Chinese Han population. Genet Test Mol Biomarkers 2013;17:407-11.
55. Zakieh A, Simin H, Forousan S, Manoochehr T. Polymorphic CT dinucleotide repeat in the GATA3 gene and risk of breast cancer in Iranian women. Med Oncol 2013;30:504.
56. Zheng Y, Huo D, Zhang J, Yoshimatsu TF, Niu Q, Olopade OI. Microsatellites in the estrogen receptor (ESR1, ESR2) and androgen receptor (AR) genes and breast cancer risk in African American and Nigerian women. PLoS One 2012;7:e40494.
57. Xie GB, Xu P, Che YN, Xia YJ, Cao YX, Wang WJ, et al. Microsatellite polymorphism in the fibrillin 3 gene and susceptibility to PCOS: a case-control study and meta-analysis. Reprod Biomed Online 2013;26:168-74.
58. Belguith-Maalej S, Kallel R, Mnif M, Abid M, Ayadi H, Kacem HH. Association of intronic repetition of SLC26A4 gene with Hashimoto thyroiditis disease. Genet Res (Camb) 2013;95:38-44.
59. Sjakste T, Eglite J, Sochnevs A, Marga M, Pirags V, Collan Y, et al. Microsatellite genotyping of chromosome 14q13.2-14q13 in the vicinity of proteasomal gene PSMA6 and association with Graves' disease in the Latvian population. Immunogenetics 2004;56:238-43.
60. Sjakste T, Trapina I, Rumba-Rozenfelde I, Lunin R, Sugoka O, Sjakste N. 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 2010;29:543-51.
61. Silva GA, Santos MP, Mota-Passos I, Boechat AL, Malheiro A, Naveca FG, et al. IFN-γ +875 microsatellite polymorphism as a potential protection marker for leprosy patients from Amazonas state, Brazil. Cytokine 2012;60:493-7.
62. Hijikata M, Shojima J, Matsushita I, Tokunaga K, Ohashi J, Hang NT, et al. Association of IFNGR2 gene polymorphisms with pulmonary tuberculosis among the Vietnamese. Hum Genet 2012;131:675-82.
63. Rodriguez-Fontenla C, Carr A, Gomez-Reino JJ, Tsezou A, Loughlin J, Gonzalez A. Association of a BMP5 microsatellite with knee osteoarthritis: case-control study. Arthritis Res Ther 2012;14:R257.
64. Michalova E, Vojtesek B, Hrstka R. Impaired pre-mRNA processing and altered architecture of 3' untranslated regions contribute to the development of human disorders. Int J Mol Sci 2013;14:15681-94.
65. Akhter Q, Masood A, Ashraf R, Majid S, Rasool S, Khan T, et al. Polymorphisms in the 3'UTR of the human leptin gene and their role in hypertension. Mol Med Rep 2012;5:1058-62.
66. Kersting C, Agelopoulos K, Schmidt H, Korsching E, August C, Gosheger G, et al. Biological importance of a polymorphic CA sequence within intron 1 of the epidermal growth factor receptor gene (EGFR) in high grade central osteosarcomas. Genes Chromosomes Cancer 2008;47:657-64.
67. Gupta R, Prakash S, Parveen F, Agrawal S. Association of CTLA-4 and TNF-α polymorphism with recurrent miscarriage among North Indian women. Cytokine 2012;60:456-62.
68. Raje M, Botre C, Ashma R. Genetic epidemiology of osteoporosis across four microsatellite markers near the VDR gene. Int J Mol Epidemiol Genet 2013;4:101-8.
69. Zeegers MP, Nekeman D, Khan HS, van Dijk BA, Goldbohm RA, Schalken J, et al. Prostate cancer susceptibility genes on 8p21-23 in a Dutch population. Prostate Cancer Prostatic Dis 2013;16:248-53.
70. Ota M, Ito T, Umemura T, Katsuyama Y, Yoshizawa K, Hamano H, et al. Polymorphism in the KCNA3 gene is associated with susceptibility to autoimmune pancreatitis in the Japanese population. Dis Markers 2011;31:223-9.

Received 13 November 2013, accepted 30 December 2013

4. DISCUSSION

In the current study we have investigated genetic diversity 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 (Publication I). The case/control scheme was used for association study with bronchial asthma in LV and TW populations (Publication III). Above listed loci and *PSMB5* (rs11543947) locus were discovered on association with JIA subtypes (Publication II), but *PSMB5* (rs11543947) and *PSMC6* (rs2295826 and rs2295827) on association with obesity in Latvians (Publication IV).

4.1. The rs11543947 (*PSMB5*) gene association analysis with autoimmune and metabolic diseases in Latvian population

Several loci of Chromosome 14 potentially responsible for susceptibility to autoimmune diseases have been identified in different human populations. A significant association was found between genetic variants of the *PSMD9* gene and obesity-associated phenotypes in Italians (Gragnoli et al., 2013). Earlier It had been reported that the rs2230087 polymorphism of the *PSMB5* gene (3'-UTR gene region) is associated with T2DM (Kim et al., 2008). In our study (Publication II, VI) we were interested in the rs11543947 of the same gene locating in exon 1 or intron 1 (*PSMB5* c.70C>T or *PSMB5* c.-112+300C>T, respectively) depending of transcript variant.

All loci we have studied here belong to the non-coding regions of corresponding genes and nucleotide substitutions potentially may influence the gene expression through allele specific targeting of different regulatory elements. Among the allele-specific targets described in current study (Publication I, II) several sites showed affinity to transcription factors and splicing signals implicated previously in immunity and lung pathology.

The presence of a minor allele at the rs11543947 locus creates a binding site to the PLU-1/JARID1B nuclear protein (Fig. 4.1.1.), which is upregulated in breast cancers, belongs to the ARID family of DNA binding proteins and has strong transcriptional repression activity. (Scibetta et al., 2007). Sequence having comonnon allele at the rs11543947 site can potentially bind GLIF/GLIS2.01 proteins responsible for the CNS development, negative regulation of cell proliferation, neurogenesis (Xu et al., 2011).

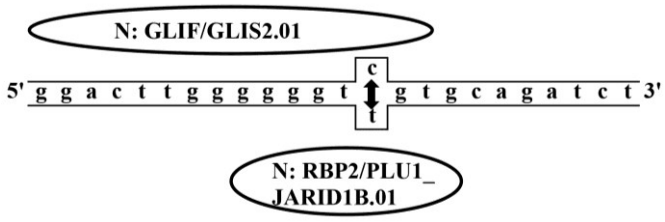


Figure 4.1.1. Consequences of the rs11543947 nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMB5* gene. Positive and negative DNA strands are indicated by capital letters P and N respectively. The transcription factors family and matrix names are separated by symbol of division and given according to MatInspector, Release 7.4 online tool at www.genomatix.de/: N: GLIF/GLIS2.01 GLI zinc finger family element; N: RBP2/PLU1_JARID1B.01 retinoblastoma-binding proteins with demethylase activity.

The rs11543947 heterozygous genotype manifested a modest association with obesity (OB) risk in total group and in children with familial OB. A rare T allele was observed more frequently ($P < 0.05$) in both these groups than in controls (Publication VI). In Latvians this SNP did not show any association with JIA (Publication II) and was identified as diseases neutral with multiple sclerosis (Kalnina et al., 2014).

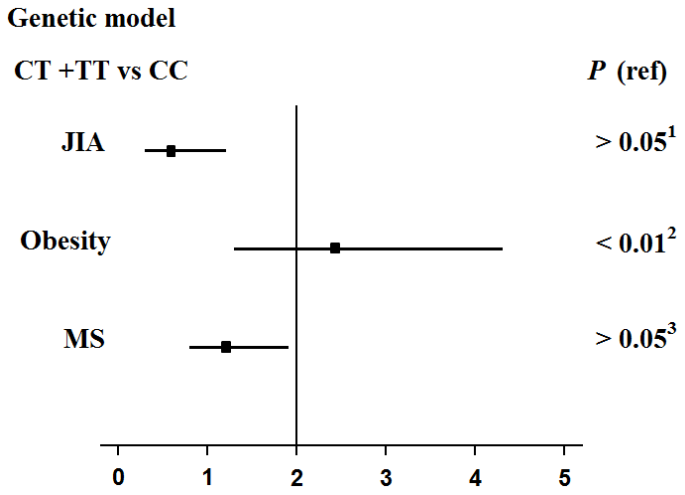


Figure 4.1.2. The forest plot of odds ratios (OR) and associated confidence intervals [95% CI] on rs11543947 locus association analysis with diseases in Latvian population. References on [1]Publication II; [2]Publication III; [3]Kalnina et al., 2014.

4.2. The rs2277460 (*PSMA6*) gene population analysis and association with autoimmune and metabolic diseases in Latvian population.

The proteasome gene, *PSMA6*, codes for a 246 residue protein called α1. This protein is structurally important in forming the outer α rings of the 20S core proteasome. The α1 protein function is also likely to be modulated by posttranslational modifications including phosphorylation, glycosylation, and lysine acetylation (Wang et al., 2013; Choudhary et al., 2009). The location of the *PSMA6* gene occurs in a region containing microsatellites that have been implicated in coronary artery disease (CAD) (Alsmadi O et al., 2009), type 2 diabetes mellitus (T2DM) (Sjakste et al., 2007) and Grave's disease (Sjakste et al., 2004a).

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. Transversion C → A appears to be rather recent mutation which happened and was fixed in Europeans in historical time and appears to be functionally neutral. Alleles' and genotypes' frequencies observed in LV and LT are similar to reported previously for Latvia (Sjakste et al., 2007; Trapina et al., 2009). Rare allele A not detected previously in Asian and African HapMap populations was identified in Taiwanese (Publication I). 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.

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 (4.2.1). The hnRNP A1 has been shown to play a key role in many human pathologies including lung cancer and response to viral pathogens (Jean-Philippe et al., 2013; Guo et al., 2013). It may influence protein-protein interactions including those with participation of NF-kB (Jean-Philippe et al., 2013) playing in turn a significant role in the asthma development and progression (Birrell et al., 2005; Edwards et al., 2009; Gagliardo et al., 2003). Additionally, it is involved in crosstalk with ubiquitin proteasome system at different levels of NF-kB and other regulatory proteins signalling pathways (Wu et al., 2013).

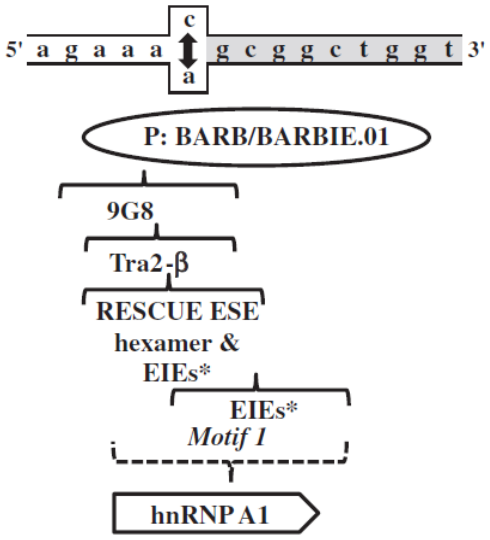


Figure 4.2.1 Consequences of the rs2277460 nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMA6* gene.

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.

The rs2277460 locus did not show any association with obesity in Latvian children. (Kupca et al., 2013), but has been detected as JIA susceptible. This conclusion is based on the results of the subtype- and sex-specific JIA association with rare allele A, heterozygous SLG, 4-LG6 heterozygous at the rs2277460, and 4-LH6 haplotype both having the rs2277460 rare allele in its structure (Publication II).

Some loci of susceptibility may be shared among many autoimmune and other immune-mediated diseases (Lee et al., 2011; Thompson et al., 2010). Earlier the genetic pleiotropic effect, as frequent phenomenon in human complex traits and diseases (Sivakumaran et al.,

2011) has been reported for asthma and obesity (Hallstrand et al., 2005; Murphy et al., 2009) and for asthma and juvenile rheumatoid arthritis (Lee et al., 2014; Ramirez-Bello et al., 2013). Similarly, rs2277460 associated with asthma in our current study (Publication III), previously have been found to be susceptible in Latvians with juvenile idiopathic arthritis (Trapina et al., 2009).

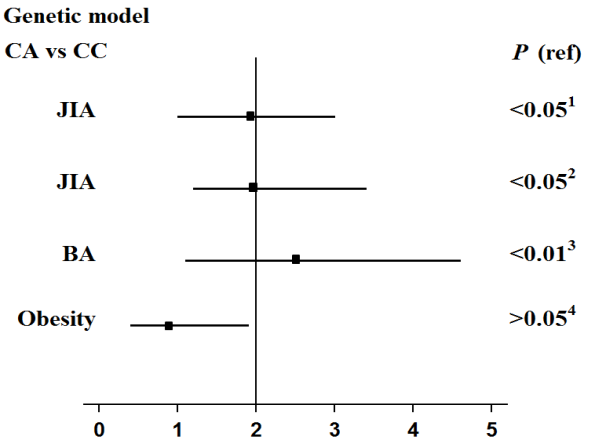


Figure 4.2.2. The forest plot of odds ratios (OR) and associated confidence intervals [95% CI] on rs2277460 locus association analysis with diseases in Latvian population. References on [1]Trapina et al., 2009; [2]Publication II; [3]Publication III; [4]Kupca et al., 2013.

On other side identification of the genetic risk factors for asthma is complicated by potential interaction of genes, metabolic pathways, genotype with sex and environment; most of the reported asthma genes were not replicated across populations (Leung et al., 2014; Loisel et al., 2011). In Taiwanese, the rs2277460 appears to be asthma neutral and resting loci showed the disease susceptibility only in females (Publication III).

4.3. The rs1048990 (*PSMA6*) gene population analysis and association with autoimmune and metabolic diseases in Latvian population

The rs1048990 SNP located in 8 nucleotides distance upstream the *PSMA6* initiation codon and four nucleotides upstream the Kozak consensus (Kozak et al., 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 (Publication I, II; Fig. 4.3.1).

Ancestral rs1048990 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 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 et al., 2012) and potentially should be involved in human adaptation.

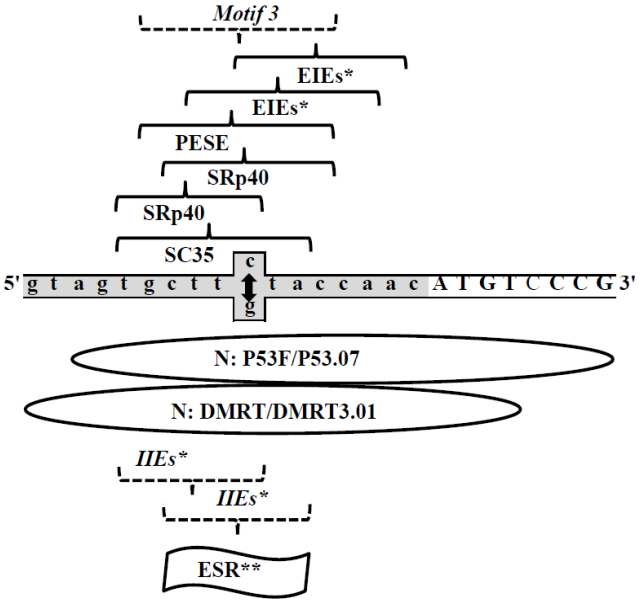


Figure 4.3.1. Consequences of the rs1048990 nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMA6* gene. Sequence features and other marks are given as described in Fig. 5.1.2.1. P53F/P53.07. tumour suppressor p53; DMRT/DMRT3.01. double sex and mab-3 related TF 3.

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 (Publication I) 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 (Publication I; Trapina et al., 2009), British (Bennett et al., 2008; Freilinger et al., 2009), 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. The rs1048990 SNP was found to occur in frequencies consistent with HWE in LV and LT groups of the current study (Publication I) and in vast majority of populations of published studies. 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 (Publication I) did not show significant deviation from neutral model of evolution.

The rs1048990 locus susceptibility reported for several pathologies (Table 3) appears to have ethnospecific character (Wang et al., 2013). We can supplement these data by results of current study. Summarizing the genotyping results obtained by different teams for association the rs1048990 with cardiovascular diseases (Alsmadi et al.,, 2009; Banerjee et al., 2008; Barbieri et al., 2008; Bennett et al., 2008; Freilinger et al., 2009; Goncharov et al., 2009; Heckman et al., 2013; Ikeda et al., 2012), T2DM (Barbieri et al., 2008; Sjakte et al., 2007), we suggested a potential of this loci to influence JIA susceptibility in Latvians. However, we did not find any association between the rs1048990 polymorphism and JIA. Similarly, locus did not show any association with obesity in Latvian children (Kupca et al., 2013).

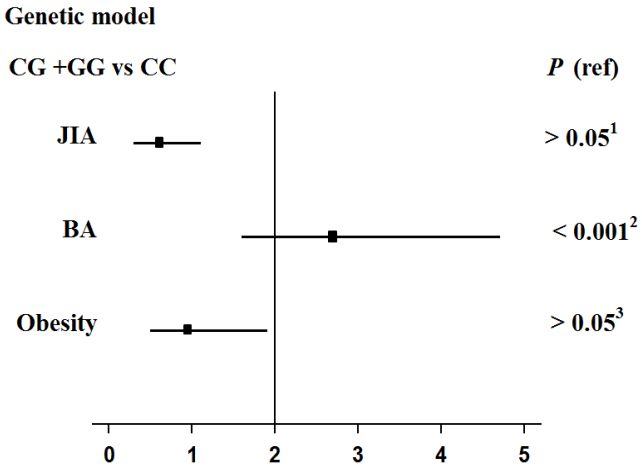


Figure 4.3.2 The forest plot of odds ratios (OR) and associated confidence intervals [95%CI] on rs1048990 locus association analysis with diseases in Latvian population. References on [1]Publication II; [2]Publication III; [3]Kupca et al., 2013.

In Latvians this locus showed the asthma main effect for rare alleles and heterozygous genotypes. In Taiwanese the asthma risk effect was observed for the rs1048990 GG rare

homozygous genotypes in female group. 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.

4.4. The rs2295826 and rs2295827 (*PSMC6*) gene population analysis and association with autoimmune and metabolic diseases in Latvian population

The rs2295826 and rs2295827 loci locate in close vicinity to each other in intron 1 of the *PSMC6* gene (*PSMC6* c.86-104A>G and *PSMC6* c.86-46C>T respectively) 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 (Fig. 4.4.1.a) 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.

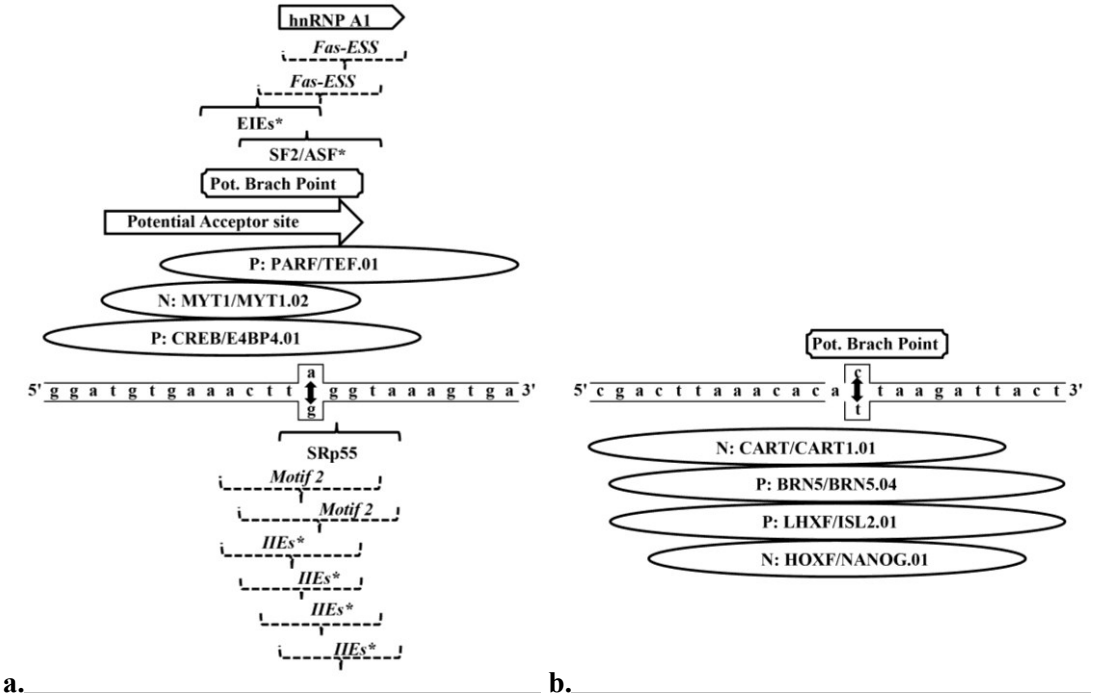


Figure 4.4.1. Consequences of the rs2295826 (a.) and rs2295827 (b.) nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMC6* gene. Sequence features and other marks are given as described in Fig. 5.1.2.1. CREB/E4BP4.01–E4BP4, bZIP domain, transcription repressor; 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.

The rs2295826 minor allele G generates splicing silencers mostly. Additional branch point is predicted for sequence encompassing the rs2295827 major allele C (Fig. 4.4.1.b). 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).

Rare alleles of these loci and their risk SLGs showed JIoA subtype-specific association by themselves and as component of risk 4-LG5 genotype and risk 4-LH4 haplotype (Publication II). In our studies have found an association between rs2295826 and rs2295827 and obesity (Publication IV); bronchial asthma (Publication III) in Latvian population (Fig. 4.4.2.). The

rs2295826 and rs2295827 heterozygotes showed asthma protective effect in Taiwanese (Publication III)

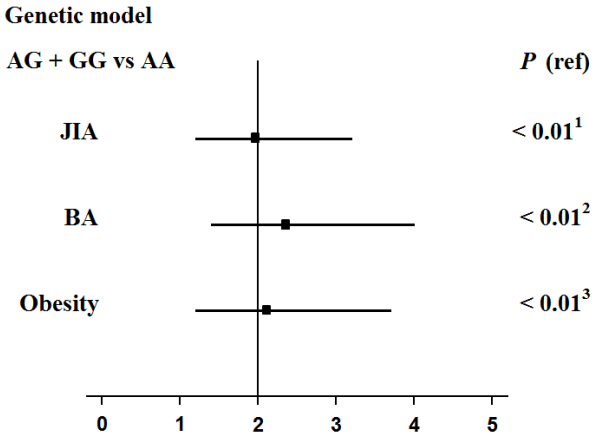


Figure 4.4.2. The forest plot of odds ratios (OR) and associated confidence intervals [95% CI] on rs2295826&rs2295826 locus association analysis with diseases in Latvian population. References on [1]Publication II; [2]Publication III; [3]Publication IV.

4.5. The rs2348071 (*PSMA3*) gene population analysis and association with autoimmune and metabolic diseases in Latvian population.

The rs2348071 locus belongs to intron 7 of the *PSMA3* gene (*PSMA3* c.543+138G>A or c.522+138G>A depending of transcript variant, 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 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 et al., 2003; Phillips et al., 2000; She et al., 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.

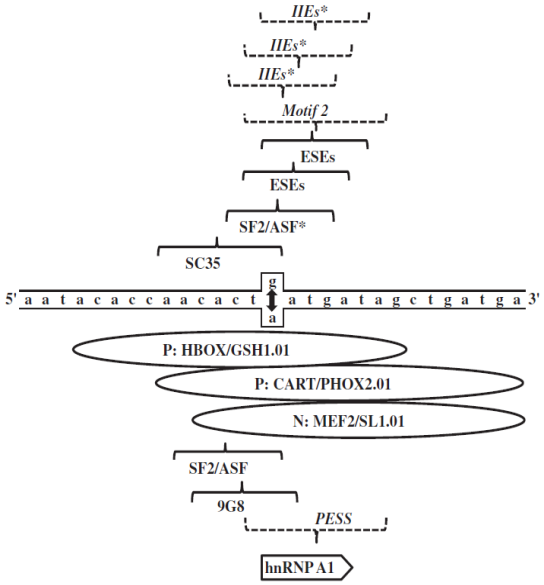


Figure 4.5.1. Consequences of the rs2348071 nucleotide substitutions on functional potential of corresponding genomic regions of the *PSMA3* gene. Sequence features and other marks are given as described in Fig. 5.1.2.1. HBOX/GSH1.01–Homeobox TF Gsh-1; Cart/PHOX2.01–Phox2a (ARIX) and Phox2b of cartilage homeoproteins family; MEF2/SL1.01–member of the RSRF related to serum response factors.

In Taiwanese the rs2348071 being in Latvian patients in HWE, significantly (P <0.001) deviated from equilibrium, and asthma risk effect was observed for rare alleles homozygotes and rare alleles. In Latvians the heterozygotes and rare alleles were found to be associated with BA (Publication III); for heterozygous genotypes was shown risk effect for JIA (Publication II). Interestingly, the rs2348071 heterozygotes were implicated previously as an obesity risk factor in Latvian children with a family history of obesity (Kupca, et al., 2013)

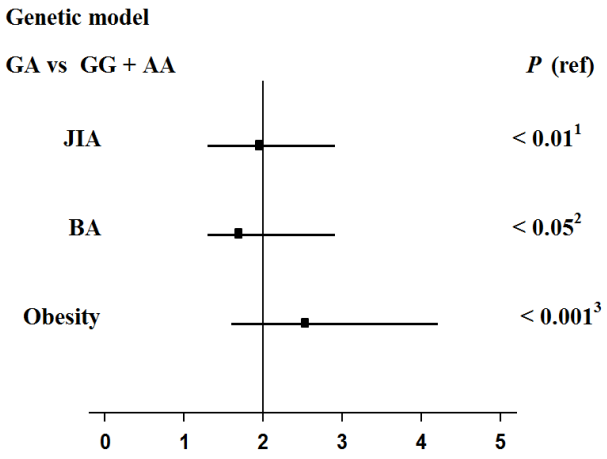


Figure 4.5.2. The forest plot of odds ratios (OR) and associated confidence intervals [95% CI] on rs2348071 locus association analysis with diseases in Latvian population. References on [1] Publication II; [2]Publication III; [3]Kupca et al., 2013.

4.6. Comparative analysis of polymorphisms

The non-additivity of diseases association in sex subgroups

Several associations revealed in our study showed non-additivity between sexes. In BA association study (Publication III) in Latvians the rs2277460 and rs2348071 were associated with asthma in females, and the rs2295826 and rs2295827 were disease susceptible in males; in Taiwanese all asthma susceptible loci were limited to females. Rare alleles and heterozygous genotypes of the rs2295826 and rs2295827 showed female-specific asthma protective effect. Sex-specific associations with the disease have been recently reported for SNPs of several genes-candidates including the *IFNG* (Loisel 2011), *IL17F* (Qian, 2012), *TSLP* (Hunninghake, 2010), *VDR* (Raby, 2004), and *KCNB1* (Seibold, 2008) genes.

In JIA association study (Publication II) in Latvians in both the control and case groups, risk genotypes were more frequent in males for the rs2277460 locus and in females for the rs2295826 and rs2295827 loci; the rs2348071 heterozygous risk genotype was more frequent in females than in males in JIpA patients.

In the current association study an interaction between sexes and familial obesity was revealed (Publication IV). The rs11543947 risk CT genotype and rare T allele were found more frequently in males with family obesity and in females without family history. The rs2295826 and rs2295827 heterozygous genotypes and minor alleles were associated with all OB subgroups in males. Sex specific differences in incidence and severity are also well known features for epidemiology of obesity. The sex-influenced association of obesity with genetic variations at the *LYPLAL1* locus, which encodes a lipase/esterase expressed in adipose tissue was suggested (Benjamin et al., 2011). Seven new loci exhibited marked sexual dimorphism with a stronger effect on weight-hip ratio in women than men (Heid et al., 2010).

Risk/protective multi-loci genotypes

It is important to note that strength of revealed association with the disease was much stronger for combination of several risk SLGs than for any individual risk SLG.

The combination of the rs2348071 (*PSMA3*) and rs2295826 and/or rs2295827 *(PSMC6)* risk genotypes represents the genetic module highly associated with both JIoA and JIpA and JIA female phenotype and plasma proteasome level in JIoA females (Fig 4.6.1.); combination of the rs2348071 and rs2277460 risk genotypes (4-LG6) represents the genetic module presumably associated with JIpA and male phenotype (Publication II).

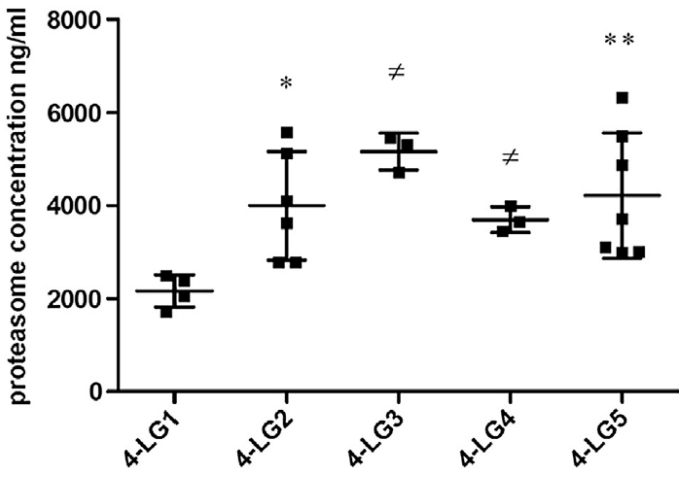


Figure 4.6.1. Plasma proteasome level in juvenile idiopathic oligoarthritis (JIoA) patients. Carriers of different four-locus genotypes (4-LGs), 4-LG1, 4-LG2, 4-LG3, 4-LG4, and 4-LG5. * $P < 0.05$, ** $P < 0.001$, # $P > 0.05$.

Earlier, circulating proteasomes were suggested as markers in autoimmune diseases (Egeger et al., 2002). Concentration of circulating proteasomes was shown to be substantially elevated in patients with rheumatoid (Egeger et al., 2002) and psoriatic (Henry et al., 2011) arthritis. The 20S proteasome has been identified as a target of the humoral autoreactive immune response in patients with systemic inflammatory diseases including autoimmune myositis (Feist et al., 1996), primary Sjögren's syndrome (Feist et al., 1999), dilated cardiomyopathy (Voigt et al., 2010), systemic lupus erythematosus (Feist et al., 1996; Arribas et al., 1991; Colmegna et al., 2008), multiple sclerosis (Fissolo et al., 2008), and psoriatic arthritis (Colmegna et al., 2008). The proteasomal inhibitor MG132 has been reported to reduce the severity of arthritis and reverse the pain behaviour in the arthritic rat models (Ahmed et al, 2010). To our knowledge, plasma levels of factors within the UPS have not been yet evaluated in JIA, and in Publication II was reported data on that for the first time.

Homozygotes on common alleles for all loci were observed more frequently in controls than obese patients (more than 80 % and less than 70 %, respectively). The rs11543947/rs2295826-rs2295827 three locus genotype heterozygous at all the loci involved was found about four times more frequently in OB patients than in controls and showed to be OB susceptible (Publication IV).

The identified asthma risk and protective single- and multi-locus genetic variants vary widely among Latvian and Taiwan populations. All variants of multi locus genotypes being homozygous for the alleles common in Latvians showed the asthma risk effect in Taiwanese females (OR = 2.911 [1.327–6.387]). Multi-loci heterozygous genotypes rs2277460/rs2348071, rs1048990/rs2348071, rs1048990/rs2295826/rs2295827 and

rs2295826/rs2295827/rs2348071 showed risk effect for bronchial asthma in Latvians, were neutral in Taiwanese. In contrast, the rs2295826/rs2295827/rs2348071 genotype of AG/CT/AA configuration being neutral in Latvians, showed the protective effect in Taiwanese females (Publication III).

Risk/protective haplotypes

Other recent experimental data suggest that haplotypes are more predictive than individual SNPs at determining risk factors for complex diseases (Kang et al., 2011). The strong association with healthy phenotype was repealed in association study between JIA and genetic variants in the *PSMA6/PSMC6/PSMA3* gene cluster, represented multi-loci haplotype C/C/G having no risk SLGs in its composition.

The T/G-T haplotype represented by rare alleles of the *PSMB5/PSMC6* genes studied loci was more frequent ($P = 0.0001$) in obese children when compared to controls (Publication IV). Haplotype diversity was higher in Latvian BA patients than in the population and did not differ between the cases and population in Taiwanese. The most frequent in Latvians the Hap1 (CCACG) showed male specific asthma protective effect. Minor Hap9 (CCGTA) was strongly ($P < 0.0001$) associated with the asthma phenotype in Latvians.

The most important finding of our study is that CGACG haplotypes with included rare alleles of rs1048990 and rs2348071 loci is a strong ($P < 0.0001$) asthma risk factor for Latvians and Taiwanese patients (Publication III).

The difference between human populations in asthma genetics is a well-known phenomenon described for a number of asthma susceptible loci having mainly ethnos specific differences in genetic diversity (Leung et al., 2014).

The Latvian and Taiwanese populations significantly differ in genetic diversity of current study loci (Publication 1). This suggests involvement of these loci in the processes of evolutional and/or geographical adaptation to environment and a potential for allele substitutions to have different ethnic specific influence on the human health and population morbidity (Publication 1).

5. CONCLUSIONS

1. Distribution of *PSMA6 (rs2277460), PSMC6 (rs2295826 and rs2295827)* and *PSMA3 (rs2348071)* proteasomal genes variations in Latvian, Lithuanian and Taiwanese populations reveals that Latvian and Lithuanian populations are significantly differentiated by each the rs1048990 ($P < 0.01$), rs2295826 and rs2295827 ($P < 0.05$) locus. Both LV, LT 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 ($P <0.05$); and exhibited similarities with LT.
2. Comparison of general and sex-specific association between the *PSMA6, PSMC6* and *PSMA3* proteasomal genes variations and childhood asthma in Latvians and Taiwanese revealed that haplotype CGACG included rare alleles at rs1048990 (*PSMA6*) and rs2348071 (*PSMA3*) is associated with the diseases ($P < 0.0001$) in Latvians and Taiwanese. All loci heterozygous genotypes and haplotype CCGTA included rare alleles at *PSMC6* SNPs and rs2348071 (*PSMA3*) are asthma risk factors in Latvians; rs1048990 and rs2348071 GG homozygotes and rs2295826 and rs2295827 heterozygotes showed asthma risk and protective effect in Taiwanese females respectively. Multi locus genotypes homozygous for alleles being common in Latvian population were identified as protective in Latvians.
3. Variations at the *PSMA6* (rs2277460), *PSMC6* (rs2295826 and rs2295827), and *PSMA3* (rs2348071) loci contribute to JIA susceptibility; combination of the rs2348071 and rs2295826 and/or rs2295827 risk genotypes (4-LG5) represents the genetic module highly associated with both JIoA and JIpA and JIA female phenotype ($P < 0.001$); combination of the rs2348071 and rs2277460 risk genotypes (4-LG6) represents the genetic module presumably associated with JIpA and male phenotype; the PSMA6/PSMC6/PSMA3 genetic variants and multiloci genetic modules could be suggested as JIA subtype- and sex-specific risk factors.
4. Plasma proteasome level was found to be significantly higher in females having four-locus risk genotypes compared to protective against JIA genotypes ($P < 0.001$).
5. Heterozygous genotype at rs11543947 (*PSMB5*) manifested association with the disease ($P < 0.01$) in total diseases group and in patients with family history of obesity. This genotype was observed more frequently ($P < 0.05$) in males with family obesity and in females without family history .The heterozygotes at rs2295826 and rs2295827 (*PSMC6*) showed nominal association ($P < 0.05$) in OB group, in males and in patients with family history. The rs11543947/rs2295826/rs2295827 multi locus genotype heterozygous at all the studied

loci and the haplotype represented by the rare alleles were more frequent in obese children when compared to controls ($P < 0.001$ and $P < 0.0001$ respectively).

6. *In silico* analysis reveals that that the nucleotide substitutions we have studied modify transcription factor binding sites and miRNAs, this can significantly modulate the transcription of related genes and gene network in response to the inflammation and other environmental stimuli and influence the diseases susceptibility.

REFERENCES

1. Abecasis G, Altshuler D, Auton A, Brooks L, Durbin R, Gibbs R, et al. A map of human genome variation from population-scale sequencing. Nature. 2010; 467:1061–73.
2. Adaikalakoteswari A, Rema M, Mohan V, Balasubramanyam M Oxidative DNA damage and augmentation of poly(ADP-ribose) polymerase/nuclear factor-kappa B signaling in patients with type 2 diabetes and microangiopathy. Int J Biochem Cell Biol. 2007; 39: 1673–84.
3. Adzhubei I, Schmidt S, Peshkin L, Ramensky V, Gerasimova A, Bork P, et. al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7: 248–49.
4. Agarwal A, Xing C, DeMartino G, Mizrachi D, Hernandez M, Sousa A, et al. *PSMB8* encoding the β5i proteasome subunit is mutated in joint contractures, muscle atrophy, microcytic anemia, and panniculitis–induced lipodystrophy syndrome. Am J Hum Genet. 2010; 87(6):866–72.
5. Aggarwal A, Sharma A, Bhatnagar A. Bi(o)communications among peripheral blood fractions: a focus on NK and NKT cell biology in rheumatoid arthritis. Autoimmunity. 2013; 46: 238–50.
6. Ahmed A, Li J, Ahmed M, Hua L, Yakovleva T, Ossipov M, et al. Attenuation of pain and inflammation in adjuvant-induced arthritis by the proteasome inhibitor MG132. Arthritis Rheum. 2010; 62:2160−9.
7. Albright J, Romero J, Saini V, Sixt S, Bird M, Kovacs E, et al. Proteasomes in human bronchoalveolar lavage fluid after burn and inhalation injury. J Burn Care Res. 2009; 30:948−56.
8. Alsmadi O, Muiya P, Khalak H, Al-Saud H, Meyer BF, Al-Mohanna F, et al. Haplotypes encompassing the *KIAA0391* and *PSMA6* gene cluster confer a genetic link for myocardial infarction and coronary artery disease. Annals of human genetics. 2009; 73(Pt 5):475–83.
9. Amos C, Gorlov I, Dong Q, Wu X, Zhang H, Lu E, et al. Nicotinic acetylcholine receptor region on chromosome 15q25 and lung cancer risk among African Americans: a case-control study. J Natl Cancer Inst. 2010; 102(15):1199–205.
10. Anna von Mikecz. The nuclear ubiquitin-proteasome system. J Cell Sci. 2006; 119(10):1977–84.
11. Anzilotti C, Pratesi F, Tommasi C, Migliorini P. Peptidylarginine deiminase 4 and citrullination in health and disease. Autoimmun Rev. 2010; 9:158–60.
12. Arribas J, Luz Rodríguez M, Alvarez-Do Forno R, Castaño JG. Autoantibodies against the multicatalytic proteinase in patients with systemic lupus erythematosus. The Journal of Experimental Medicine, 1991; 173:423–27.
13. Arima K, Kinoshita A,.Mishima H,.Kanazawa H,.Kaneko N, Mizushima T, et al. Proteasome assembly defect due to a proteasome subunit beta type 8 (PSMB8) mutation causes the autoinflammatory disorder, Nakajo-Nishimura syndrome. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108(36):14914–19.
14. Arya R, Hare E, Del Rincon I, Jenkinson C, Duggirala R, Almasy L, et al. Effects of covariates and interactions on a genome-wide association analysis of rheumatoid arthritis. BMC Proc 2009; 3:S84.
15. Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene) Genome-wide association study identifies new multiple sclerosis susceptibility on chromosomes 12 and 20. Nat Genet. 2009; 41:824–28.

16. Bachmann H, Novotny J, Sixt S, Liebisch P, Frey U, Duhrsen U, et al. The G-Allele of the PSMA6-8C>G polymorphism is associated with poor outcome in multiple myeloma independently of circulating proteasome serum levels. Eur J haematol 2010; 85(2):108–13.
17. Bader H, Hsu T. Systemic VHL gene functions and the VHL disease. FEBS Lett. 2012; 586(11):1562–69.
18. Banerjee I, Gupta V, Ahmed T, Faizaan M, Agarwal P, Ganesh S. Inflammatory system gene polymorphism and the risk of stroke: a case–control study in an Indian population. Brain research bulletin. 2008; 75(1):158–65.
19. Baptista M, Duarte C, Maciel P. Role of the ubiquitin–proteasome system in nervous system function and disease: using C. elegans as a dissecting tool. Cell Mol Life Sci. 2012; 69(16):2691–715.
20. Barbieri M, Marfella R, Rizzo M, Boccardi V, Siniscalchi M, Schiattarella C, et al. The –8 UTR C/G polymorphism of PSMA6 gene is associated with susceptibility to myocardial infarction in type 2 diabetic patients. Atherosclerosis. 2008; 201:117–23.
21. Bassuny W, Ihara K, Sasaki Y, Kuromaru R, Kohno H, Matsuura N, et al. A functional polymorphism in the promoter/ enhancer region of the FOXP3/Scurfin gene associated with type 1 diabetes. Immunogenetics. 2003; 55:149–56.
22. Belin A, Westerlund M. Parkinson's disease: a genetic perspective. FEBS J. 2008; 275(7):1377–83.
23. Bedford L, Hay D, Devoy A, Paine S, Powe D, Seth R, et al. Depletion of 26S proteasomes in mouse brain neurons causes neurodegeneration and lewy–like inclusions resembling human pale bodies. Journal of Neuroscience, 2008; 28: 8189–98.
24. Benanti J. Coordination of cell growth and division by the ubiquitin-proteasome system. Semin Cell Dev Biol. 2012; 23(5):492–98.
25. Bendotti C, Marino M, Cheroni C, Fontana E, Crippa V, Poletti A, et al. Dysfunction of constitutive and inducible ubiquitin-proteasome system in amyotrophic lateral sclerosis: implication for protein aggregation and immune response. Prog Neurobiol. 2012; 97(2):101–26.
26. Bendikov-Bar I, Horowitz M. Gaucher disease paradigm: from ERAD to comorbidity. Hum Mutat. 2012; 33(10):1398–407.
27. Bennett D, Xu P, Clarke R, Zondervan K, Parish S, Palmer A, et al. 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 2008; 16(4):480–6.
28. Bennet L, Lindblad U, Franks PW. A family history of diabetes determines poorer glycaemic control and younger age of diabetes onset in immigrants from the Middle East compared with native Swedes. Diabetes Metab. 2014; 9(10):e109506
29. Benjamin A, Suchindran S, Pearce K, Rowell J, Lien L. Gene by Sex Interaction for Measures of Obesity in the Framingham Heart Study. J Obes. 2011; 8: 329038.
30. Berto G, Camera P, Fusco C, Imarisio S, Ambrogio C, Chiarle R et al. The Down syndrome critical region protein TTC3 inhibits neuronal differentiation via RhoA and Citron kinase. J Cell Sci. 2007; 120(11):1859–67.
31. Birrell M, Hardaker E, Wong S, McCluskie K, Catley M, De Alba J, et al. Ikappa-B kinase-2 inhibitor blocks inflammation in human airway smooth muscle and a rat model of asthma. Am J Respir Crit Care Med. 2005; 172:962−71.
32. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002; 16:6–21.
33. Boyle A, Davis S, Shulha H, Meltzer P, Margulies E, Weng Z. High-resolution mapping and characterization of open chromatin across the genome. Cell. 2008; 132:311–22.

34. Boyle A, Hong E, Hariharan M, Cheng Y, Schaub M, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012; 22(9):1790-97.
35. Boulikas T. Nature of DNA sequences at the attachment regions of genes to the nuclear matrix. J Cell Biochem. 1993; 52:14–22.
36. Bo Kim K, Fonseca FN, Crews CM. Development and characterization of proteasome inhibitors. Methods Enzymol. 2005;399:585−609.
37. Brand O, Gough S. Immunogenetic mechanisms leading to thyroid autoimmunity: recent advances in identifying susceptibility genes and regions. Curr Genomics. 2011; 12: 526–41.
38. Broide D. Molecular and cellular mechanisms of allergic disease. J Allergy Clin Immunol. In practice. 2001; 108:S65−71.
39. Bruce K. Cagampang F. Epigenetic priming of the metabolic syndrome. Toxicol. Mech. Methods. 2011; 21(4):353–61.
40. Buiting K. Prader-Willi syndrome and Angelman syndrome. Am J Med Genet C Semin Med Genet. 2010; 154C(3):365–76.
41. Burgner D, Davila S, Breunis W, Ng S, Li Y, Bonnard C, et al. A genome-wide association study identifies novel and functionally related susceptibility loci for Kawasaki disease. PLoS Genet 2009; 5:e1000319.
42. Carrard G, Bulteau A, Petropoulos I, Friguet B. Impairment of proteasome structure and function in aging. International Journal of Biochemistry and Cell Biology. 2002. 34(11):1461–74,
43. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics. 2005; 21:2933−42.
44. Carvajal-Carmona L, Cazier J, Jones A, Howarth K, Broderick P, Pittman A, et al. Finemapping of colorectal cancer susceptibility loci at 8q23.3, 16q22.1 and 19q13.11: Refinement of association signals and use of in silico analysis to suggest functional variation and unexpected candidate target genes. Hum Mol Genet. 2011; 20:2879–88.
45. Cerosaletti K , Buckner J. Protein tyrosine phosphatases and type 1 diabetes: genetic and functional implications of PTPN2 and PTPN22.Rev Diabet Stud. 2012; 9:188–200.
46. Chang T, Chang C, Lee W, Lin M, Huang Y, Fan K. The roles of ubiquitin and 26S proteasome in human obesity. Metabolism. 2009; 58(11):1643–48.
47. Chen M, Rockel T, Steinweger G, Hemmerich P, Risch J, von Mikecz A. Subcellular recruitment of fibrillarin to nucleoplasmic proteasomes: implications for processing of a nucleolar autoantigen. Mol Biol Cell. 2002; 13(10):3576–87.
48. Chen M, von Mikecz A. Proteasomal processing of nuclear autoantigens in systemic autoimmunity. Autoimmun Rev. 2005a; 4(3):117–22.
49. Chen M, von Mikecz A. Xenobiotic-induced recruitment of autoantigens to nuclear proteasomes suggests a role for altered antigen processing in scleroderma. Ann N Y Acad Sci. 2005b; 1051:382–89.
50. Chen H, Sun C, Guo W, Meng R, Du H, et al. AluYb8 insertion in the MUTYH gene is related to increased 8–OHdG in genomic DNA and could be a risk factor for type 2 diabetes in a Chinese population. Mol Cell Endocrinol. 2011; 332: 301–5.
51. Chistyakov D, Savost'anov K, Turakulov R, Nosikov V. Genetic determinants of Graves disease. Mol Genet Metab 2000; 71:66e9.
52. Chin J, Schleifman E, Glazer P. Repair and recombination induced by triple helix DNA. Front Biosci. 2007; 12:4288−97.

53. Chitra S, Nalini G, Rajasekhar G. The ubiquitin proteasome system and efficacy of proteasome inhibitors in diseases. Int J Rheum Dis. 2012; 15(3):249–60.
54. Chiusolo P1, Giammarco S, Bellesi S, Metafuni E, Piccirillo N, et al. The role of MTHFR and RFC1 polymorphisms on toxicity and outcome of adult patients with hematological malignancies treated with high-dose methotrexate followed by leucovorin rescue. Cancer Chemother Pharmacol. 2012; 69: 691–96.
55. Choe J, Park S, Kim S. MTHFR polymorphisms, C677T and A1298C, are not a genetic risk factor for liver stiffness index in methotrexate–treated rheumatoid arthritis patients. Clin Exp Rheumatol. 2011; 29:144–45.
56. Chondrogianni N, Petropoulos I, Franceschi C, Friguet B,. Gonos E. Fibroblast cultures fromhealthy centenarians have an active proteasome. Experimental Gerontology. 2000; 35(6–7):721–28.
57. Chu–Ping M, Vu J, Proske R, Slaughter C, DeMartino G. Identification, purification, and characterization of a high molecular weight, ATP-dependent activator (PA700) of the 20 S proteasome. J Biol Chem. 1994; 269(5):3539–47.
58. Ciechanover A, Orian A, Schwartz A. Ubiquitin-mediated proteolysis: biological regulation via destruction. BioEssays:news and reviews in molecular, cellular and developmental biology. 2000; 22(5):442–51.
59. Ciechanover A. Intracellular Protein Degradation: From a Vague Idea through the Lysosome and the Ubiquitin-Proteasome System and onto Human Diseases and Drug Targeting. Neurodegenerative Diseases 2012; 10(1–4):7–22.
60. Clower C, Chatterjee D, Wang Z, Cantley L, Van der Heiden M, Krainer A. The alternative splicing repressors hnRNP A1/A2 and PTB influence pyruvate kinase isoform expression and cell metabolism. Proc Natl Acad Sci U S A. 2010; 107:1894−9.
61. Colmegna I, Sainz B Jr, Citera G, Maldonado-Cocco JA, Garry RF, Espinoza LR. Anti–20S proteasome antibodies in psoriatic arthritis. The Journal of Rheumatology. 2008; 35:674–76.
62. Corne´lis F, Faure´ S, Martinez M, Prud'homme J, Fritz P, Dib C, et al. New susceptibility locus for rheumatoid arthritis suggested by a genome–wide linkage study. Proc Natl Acad Sci U S A 1998; 95:10746e50.
63. Comi C, Cappellano G, Chiocchetti A, Orilieri E, Buttini S, et al. The impact of osteopontin gene variations on multiple sclerosis development and progression. Clin Dev Immunol 2012; 212893.
64. Compston A, Confavreux C, Lassmann H, et al. McAlpine's multiple sclerosis. Philadelphia: Churchill Livingstone/Elsevier 2006.
65. Costa Mdo C, Paulson H. Toward understanding Machado-Joseph disease. Prog Neurobiol. 2012; 97(2):239–57.
66. Coux O, Tanak K, Goldberg A. Structure and functions of the 20S and 26S proteasomes. Annu. Rev. Biochem. 1996; 65:801–47.
67. Cruciani F, Trombetta B, Labuda D, Modiano D, Torroni A, Costa R, et al. Genetic diversity patterns at the human clock gene period 2 are suggestive of population-specific positive selection. Eur J Hum Genet. 2008; 16(12):1526–34.
68. Dasuri K, Zhang L, Ebenezer P, Liu Y, Fernandez-Kim S, Keller J. Aging and dietary restriction alter proteasome biogenesis and composition in the brain and liver. Mechanisms of Ageing and Development. 2009; 130(11–12):777–83.
69. Day FR, Loos RJ. Developments in obesity genetics in the era of genome-wide association studies. J Nutrigenet Nutrigenomics. 2011; 4(4):222–38.

70. Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R. Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome Science. 1998; 280:1077-82
71. De Jager P, Jia X, Wang P, de Bakker L, Ottoboni L, et al. International MS Genetics Consortium. Meta-analysis of genome scans and replication identityCD6, IRF8 and TNFRF1A as new multiple sclerosis susceptibility loci. Nat genet. 2009; 41:776–82.
72. Dennissen F, Kholod N, van Leeuwen F. The ubiquitin proteasome system in neurodegenerative diseases: culprit, accomplice or victim? Prog Neurobiol. 2012; 96(2):190–207.
73. Desmet F, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009; 37:e67.
74. Dino Rockel T, von Mikecz A. Proteasome-dependent processing of nuclear proteins is correlated with their subnuclear localization. J Struct Biol. 2002; 140(1-3):189–99.
75. Dominguez-Villar M, Baecher-Allan CM, Hafler DA Identification of T helper type 1-like, Foxp3+ regulatory T cells in human autoimmune disease. Nat Med. 2011; 17:673–75.
76. Dozmorov M, WuW, Chakrabarty K, Booth J, Hurst R, Coggeshall K, et al. 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. 2009; 9:152
77. Edwards M, Bartlett N, Clarke D, Birrell M, Belvisi M, Johnston S. Targeting the NF-kappaB pathway in asthma and chronic obstructive pulmonary disease. Pharmacol Ther. 2009; 121:1−13.
78. Egerer K, Trimpert C, Feist E, Gericke C, et al. Humoral anti-proteasomal autoimmunity in dilated cardiomyopathy. Basic Res Cardiol. 2001; 105:9–18.
79. Egerer K, Kuckelkorn U, Rudolph P, Rückert J, Dörner T, Burmester G, et al. Circulating proteasomes are markers of cell damage and immunologic activity in autoimmune diseases. The Journal of Rheumatology. 2002; 29:2045–52.
80. Elliott P, Zollner T, Boehncke W. Proteasome inhibition: a new anti-inflammatory strategy. J Mol Med. 2003; 81:235–45
81. Emahazion T, Feuk L, Jobs M, Sawyer S, Fredman D, St Clair D, et al. SNP association studies in Alzheimer's disease highlight problem for complex disease analysis. Trends Genet. 2001; 17:407–13.
82. Ernst J, Kheradpour P, Mikkelsen T, Shoresh N, Ward L, Epstein C, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 2011; 473:43–49.
83. Esteller M: Epigenetics in cancer. N Engl J Med. 2008; 358:1148–59.
84. Fabunmi R, Wigley W, Thomas P, DeMartino G. Activity and regulation of the centrosome-associated proteasome. J Biol Chem. 2000; 275:409–13.
85. Fallahi G, Sabbaghian M, Khalili M, Parvaneh N, Zenker M, Rezaei N. Novel UBR1 gene mutation in a patient with typical phenotype of Johanson-Blizzard syndrome. Eur J Pediatr. 2011; 170(2):233–35.
86. Feist E, Dörner T, Kuckelkorn U, Schmidtke G, Micheel B, Hiepe F, et al. Proteasome alpha-type subunit C9 is a primary target of autoantibodies in sera of patients with myositis and systemic lupus erythematosus. The Journal of Experimental Medicine. 1996; 184:1313–18.
87. Feist E, Kuckelkorn U, Dörner T, Dönitz H, Scheffler S, Hiepe F, et al. Autoantibodies in primary Sjögren's syndrome are directed against proteasomal subunits of the alpha and beta type. Arthritis & Rheumatism. 1999; 42:697–702.

88. Fellerhoff B, Gu S, Laumbacher B, Nerlich B, Weiss A, Glas E, et al. The LMP7-K allele of the immunoproteasome exhibits reduced transcript stability and predicts high risk of colon cancer. Cancer Research. 2011; 71(23):7145–54.
89. Fenoglio C, Ridolfi E, Galimberti D, Scarpini E MicroRNAs as Active Players in the Pathogenesis of Multiple Sclerosis. Int J Mol Sci. 2012; 13:13227–39.
90. Fierabracci A. Proteasome inhibitors: a new per spective for treating autoimmune diseases. CurrentDrug Targets. 2012; 13:1665–75.
91. Fife B, Bluestone J. Control of peripheral T-cell tolerance and autoimmunity via the CTLA-4 and PD-1 pathways. Immunol Rev. 2008; 224:166–82.
92. Filippova GN, Thienes CP, Penn BH, Cho DH, Hu YJ, Moore JM, et al. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat Genet. 2001; 28:335–43.
93. Fissolo N, Kraus M, Reich M, Ayturan M, Overkleeft H, et al. Dual inhibition of proteasomal and lysosomal proteolysis ameliorates autoimmune central nervous system inflammation. Eur J Immunol 2008; 38:2401–11.
94. Flick K, Kaiser P. Protein degradation and the stress response. Semin Cell Dev Biol. 2012; 23(5):515–522.
95. Fousteri G, Liossis SN, Battaglia M. Roles of the protein tyrosine phosphatase PTPN22 in immunity and autoimmunity. Clin Immunol. 2013; 149:556–65.
96. Freilinger T, Bevan S, Ripke S, Gschwendtner A, Lichtner P, Müller-Myhsok B, et al. Genetic variation in the lymphotoxin-alpha pathway and the risk of ischemic stroke in European populations. Stroke. 2009; 40(3):970–72.
97. Furukawa K, Iioka T, Morishita M, Yamaguchi A, Shindo H, Namba H, et al. Functional domains of paired–like homeoprotein Cart1 and the relationship between dimerization and transcription activity. Genes Cells. 2002; 7:1135e47.
98. Gaczynska M, Osmulski P. Small-molecule inhibitors of proteasome activity. Methods Mol Biol. 2005; 301:3–22.
99. Gagliardo R, Chanez P, Mathieu M, Bruno A, Costanzo G, Gougat C, et al. Persistent activation of nuclear factor-kappaB signaling pathway in severe uncontrolled asthma. J Respir Crit Care Med. 2003; 168:1190−98.
100. Galli SJ, Tsai M, Piliponsky AM The development of allergic inflammation. Nature. 2008; 454:445–54.
101. George N, DeMartino, Clive A. Slaughter. The Proteasome, a Novel Protease Regulated by Multiple Mechanisms. The Journal of Biological Chemistry. 1999; 274(32):22123–26.
102. Gill G. Decoding the LIM development code. Trans Am Clin Climatol Assoc 2003; 114:179e89.
103. Glickman M, Rubin D, Coux O, Wefes I, Pfeifer G, Cjeka Z, et al. A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3.Cell. 1998; 94:615–23.
104. Goldbach-Mansky R. Immunology in clinic review series; focus on autoinflammatory diseases: update on monogenic autoinflammatory diseases: the role of interleukin (IL)-1 and an emerging role for cytokines beyond IL-1. Clin Exp Immunol. 2012; 167(3):391–404.
105. Goldberg A, Cascio P, Saric T, Rock K. The importance of the proteasome and subsequent proteolytic steps in the generation of antigenic peptides. Mol Immunol. 2002; 39:147–64.
106. Gomes A. Genetics of proteasome diseases. Scientifica (Cairo). 2013; 637629.

107. Goncharov S, Dosenko V, Haitovich N, Moĭbenko A. Allele polymorphism of genes coding proteasome subunits is associated with an enhanced risk for arterial hypertension in adolescents. Fiziol Zh. 2009; 55:3–10.
108. Gondro C, Lee S, Lee H, Porto-Neto L. Quality control for genome-wide association studies. Methods Mol Biol. 2013; 1019:129–47.
109. Gragnoli C. Proteasome modulator 9 and macrovascular pathology of T2D. Cardiovascular diabetology 2011; 10:32.
110. Grandi F, An W. Non-LTR retrotransposons and microsatellites: partners in genomic variation. Mob Genet Elements. 2013; 3:e25674.
111. Grecchi S, Mazzini G, Lisa A, Armentero M, Bergamaschi R, et al. Search for cellular stress biomarkers in lymphocytes from patients with multiple sclerosis: a pilot study. PLoS One. 2012; 7: e44935.
112. Grisham MB, Palombella VJ, Elliott PJ, et al. Inhibition of NF-κB activation in vitro and in vivo: role of 26S proteasome. Methods Enzymol. 1999; 300:345−63.
113. Groll M, Ditzel L, Lowe J, Stock D, Bochtler M, Bartunik H. Structure of 20S proteasome from yeast at 2.4 A resolution, Nature. 1997; 386:463–71.
114. Gross DS, Garrard W. Nuclease hypersensitive sites in chromatin. Annu Rev Biochem. 1988; 57:159–97.
115. Crawford G, Holt I, Whittle J, Webb B, Tai D, Davis S, et al. Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). 2006; Genome Res. 16:123–31.
116. Gulcher J. Microsatellite markers for linkage and association studies. Cold Spring Harb Protoc. 2012; 4:425–32.
117. Hafler D, Compston A, Sawcer S, Lander E, Daly M, et al. Risk allels for multiple sclerosis identified by a genomewide study. International Multiple Sclerosis Genetics Consortium. N Engl J Med. 2007; 357:851–62
118. Hakonarson H, Bjornsdottir U, Halapi E, Palsson S, Adalsteinsdottir E, Gislason D, et al. A major susceptibility gene for asthma maps to chromosome 14q24 Am J Hum Genet. 2002a; 71:483−91.
119. Hakonarson H, Halapi E. Genetic analyses in asthma: current concepts and future directions. Am J Pharmacogenomics. 2002b; 2(3):155–66.
120. Hamazaki J, Sasaki K, Kawahara H, Hisanaga S, Tanaka K, Murata S. Rpn10-mediated degradation of ubiquitinated proteins is essential for mouse development. Molecular and Cellular Biology. 2007; 27:6629–38.
121. Hansen H, Xiao Y, Rice T, Bracci P, Wrensch M, Sison J, et al. Fine mapping of chromosome 15q25.1 lung cancer susceptibility in African-Americans. Human molecular genetics 2010; 19(18):3652–61.
122. HapMap Consortium. The international HapMap Project. Nature. 2003; 426:789–96.
123. Harismendy O, Notani D, Song X, Rahim N, Tanasa B, Heintzman N, et al. 9p21 DNA variants associated with coronary artery disease impair interferon-g signalling response. Nature. 2011; 470:264–68.
124. Hawkins G, Meyers D, Bleecker E, Pack A. 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. 2008;19(1):44–9.
125. Heckman M, Soto–Ortolaza A, Diehl N, Rayaprolu S, Brott T, Wszolek Z, et al. Genetic variants associated with myocardial infarction in the PSMA6 gene and Chr9p21 are also associated with ischaemic stroke. Eur. J. Neurol. 2013; 20:300–08.

126. Heid I, Jackson A, Randall J, Winkler T, Qi L. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nat Genet. 2010; 42: 949–60.
127. Henry L, Le Gallic L, Garcin G, Coux O, Jumez N, Roger P, et al. Proteolytic activity and expression of the 20S proteasome are increased in psoriasis lesional skin. British Journal of Dermatology. 2011; 165:311–20.
128. Hershko A, Ciechanover A. The ubiquitin system. Annu Rev Biochem. 1998; 67:425–79.
129. Hesselberth J, Chen X, Zhang Z, Sabo P, Sandstrom R, Reynolds A, et al. Global mapping
130. of protein-DNA interactions in vivo by digital genomic footprinting. Nat Methods. 2009 6:283–89.
131. Hinohara K, Nakajima T, Sasaoka T, Sawabe M, Lee BS, Ban J, et al. Replication studies for the association of PSMA6 polymorphism with coronary artery disease in East Asian populations. J. Hum. Genet. 2009; 54:248–51.
132. Hipkiss R. Accumulation of altered proteins and ageing: causes and effects. Experimental Gerontology. 2006; 41(5):464–73.
133. Ho B, Olson G, Figel S, Gelman I, Cance W, Golubovskaya V. Nanog increases focal adhesion kinase (FAK) promoter activity and expression and directly binds to FAK protein to be phosphorylated. J Biol Chem. 2012; 287:18656e73.
134. Hoffman L, Pratt G, Rechsteiner M. Multiple forms of the 20 S multicatalytic and the 26 S ubiquitin/ATP-dependent proteases from rabbit reticulocyte lysate. J. Biol. Chem. 1992; 267:22362–68.
135. Hogg K, Price E, Hanna C, Robinson W. Prenatal and perinatal environmental influences on the human fetal and placental epigenome. Clin. Pharmacol. Ther. 2012; 92(6):716–26.
136. Honcharov S, DosenkoV, Khaitovych M. Allele polymorphism of genes coding proteasome subunits is associated with an enhanced risk for arterial hypertension in adolescents. 2009; Fiziol. Zh. 55:3–10.
137. Hsu S, Chu C, Tsou A, Chen S, Chen H, Hsu P, et al. miRNAMap 2.0: genomic maps of microRNAs in metazoan genomes. Nucleic Acids Res. 2008; 36:D165−9.
138. Hui J, Hung L, Heiner M, Schreiner S, Neumuller N, Reither G, et al. Intronic CA-repeat and CA-rich elements: a new class of regulators of mammalian alternative splicing. EMBO J. 2005; 24:1988–98.
139. Hunninghake G, Soto–Quirós M, Avila L, Kim H, Lasky-Su J, Rafaels N, et al. TSLP polymorphisms are associated with asthma in a sex-specific fashion. Allergy 2010; 65:1566−75.
140. Ihara Y, Morishima-Kawashima M, Nixon R. The ubiquitin-proteasome system and the autophagic-lysosomal system in Alzheimer disease. Cold Spring Harb Perspect Med. 2012; 2(8).
141. Ikeda S, Tanaka N, Arai T, Chida K, Muramatsu M, Sawabe M, et al. Polymorphisms of LTA, LGALS2, and PSMA6 genes and coronary atherosclerosis: a pathological study of 1503 consecutive autopsy cases. Atherosclerosis. 2012; 221:458–60.
142. International Multiple Sclerosis Genetics Consorciaum (IMSGC) IL12A, MPHOSPH9/CDK2AP1 and RGS1 are novel multiple sclerosis susceptibility loci. Genes Immun. 2010; 11:397–405.
143. International Multiple Sclerosis Genetics Consorciaum (IMSGC) and Welcome Trust case Control Consortium 2 (WTCCC2) Genetic risk and a primary role for cell–mediated immune mechanisms in multiple sclerosis. Nature. 2011; 476:214–19.
144. Ishibashi T. Molecular Hydrogen: New Antioxidant and Anti-inflammatory Therapy for Rheumatoid Arthritis and Related Diseases. Curr Pharm Des. 2013; 19:6375–81.

145. Jarinova O, Stewart A, Roberts R,Wells G, Lau P, Naing T, et al. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. Arterioscler Thromb Vasc Biol. 2009; 29:1671–77.
146. Jarne P, Lagoda P. Microsatellites, from molecules to populations and back. Trends Ecol Evol. 1996; 11:424–9.
147. Jean-Philippe J, Paz S, Caputi M. hnRNP A1: the Swiss army knife of gene expression. Int J Mol Sci. 2013; 14:18999−19024.
148. Jeffrey C. Genome-wide association study and meta-analysis finds over 40 loci affect risk of type 1 diabetes. Nat Genet. 2009; 41(6): 703–07.
149. Jimenez-Sanchez M, Thomson F, Zavodszky E, Rubinsztein DC. Autophagy and polyglutamine diseases. Prog Neurobiol. 2012; 97(2):67–82.
150. Johnson D, Mortazavi A, Myers R,Wold B. Genome-wide mapping of in vivo protein-DNA interactions. Science. 2007; 316:1497–502.
151. Julia A, Ballina J, Cañete JD, Balsa A, Tornero-Molina J, et al. Genome–wide association study of rheumatoid arthritis in the Spanish population: KLF12 as a risk locus for rheumatoid arthritis susceptibility. Arthritis Rheum. 2008; 58:2275–86.
152. Kalis M, Sjakste T, Sjakste N, Luthman H, Groop L. Association study between (TG) repeat polymorphism in *PSMA6* gene and type II diabetes mellitus in Botnia. Biologija (Vilnius) 2003; 2:12–4.
153. Kanazawa N. Nakajo-Nishimura syndrome: an autoinflammatory disorder showing pernio-like rashes and progressive partial lipodystrophy. Allergol Int. 2012; 61(2):197–206.
154. Kang J, Kugathasan S, Georges M, Zhao H, Cho J. Improved risk prediction forCrohn's diseasewith amulti-locus approach. HumanMolecular Genetics. 2011; 20(12):2435–42.
155. Kasowski M, Grubert F, Heffelfinger C, Hariharan M, Asabere A,Waszak SM, et al. Variation in transcription factor binding among humans. Science. 2010; 328:232–35.
156. Kasznicki J, Krupa R, Błasiak J, Drzewoski J. Association between polymorphisms of the DNA repair genes XRCC1 and hOGG1 and type 2 diabetes mellitus in the Polish population. Pol Arch Med Wewn. 2009; 119:122–28.
157. Kellermayer R. Epigenetics and the developmental origins of inflammatory bowel disease. Canadian J. Gastroenterol. 2012; 26(12):909–15.
158. Kim J, Mirkin S. The balancing act of DNA repeat expansions. Curr Opin Genet Dev. 2013; 23:280–8.
159. King DG. Evolution of simple sequence repeats as mutable sites. Adv Exp Med Biol. 2012; 769:10V25.
160. Kish-Trier and C Hill. Structural biology of the proteasome. Annual Review of Biophysics, 2013; 42:29–49.
161. Kisselev A, Garcia-Calvo M, Overkleeft H, Peterson E, Pennington M, Ploegh H, et al. The caspase-like sites of proteasomes, their substrate specificity, new inhibitors and substrates, and allosteric interactions with the trypsin-like sites. The Journal of biological chemistry. 2003; 278(38):35869–77.
162. Kitamura A, Maekawa Y, Ueehara H, Izumi K, Kawachi I, Nishizawa M, et al. A mutation in the immunoproteasome subunit *PSMB8* causes autoinflammation and lipodystrophy in humans. Clin Invest. 2011; 121:4150–60.

163. Kozak M. 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. 1997; 16(9):2482–92.
164. Kloetzel P. Antigen processing by the proteasome. Nat Rev Mol Cell Biol. 2001; 2:179–87.
165. Kofler DM, Severson CA, Mousissian N, De Jager PL, Hafler DA The CD6 multiple sclerosis susceptibility allele is associated with alterations in CD4+ T cell proliferation. J Immunology. 2011; 187:3286–91.
166. Konstantinova I, Tsimokha A, Mittenberg A. Role of proteasomes in cellular regulation. International review of cell and molecular biology. 2008; 267:59−124.
167. Kopp A. *Dmrt* genes in the development and evolution of sexual dimorphism. Trends Genet. 2012; 28(4):175–84.
168. Korman B, Kastner D, Gregersen P, Remmers E. STAT4: Genetics, Mechanisms, and Implications for Autoimmunity Review for Current Allergy and Asthma Reports. Curr Allergy Asthma Rep.2008; 8:398–403.
169. Kornek B, Storh M, Weisser R, Wallstroem E, Stefferl A, et al. Multiple sclerosis and chronic autoimmune encephalomyelitis: a comparative quantitative study of axon injury in active, inactive and remyelinated lesions. Am J Pathol. 2000; 157:267–76.
170. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. 2011; 39:D152−57.
171. Kristiansen OP, Larsen ZM, Pociot F. (2000) *CTLA-4* in autoimmune diseases-a general susceptibility gene to autoimmunity? Genes Immun. 1:170–184.
172. Kupca S, Sjakste T, Paramonova N, Sugoka O, Rinkuza I, Trapina I, et al. Association of obesity with proteasomal gene polymorphisms in children. Journal of obesity. 2013; 638154.
173. Kurkó J, Besenyei T, Laki J, Glant TT, Mikecz K, Szekanecz Z. Genetics of rheumatoid arthritis - a comprehensive review. Clin Rev Allergy Immunol. 2013; 45:170–79.
174. Landers J, Melki J, Meininger V, Glass J, van den Berg L, van Es M, et al. Reduced expression of the Kinesin-Associated Protein 3 (*KIFAP3*) gene increases survival in sporadic amyotrophic lateral sclerosis. Proc Natl Acad Sci. 2009; 106:9004–09.
175. Lee J, Jeon M, Seo Y, Lee Y, Ko J, Tsujimoto Y, et al. CA repeats in the 3f-untranslated region of bcl-2 mRNA mediate constitutive decay of bcl-2 mRNA. J Biol Chem. 2004; 279:42758–64.
176. Lee S, Lee E, Shin E, Lee J, Cho S, Min K, et al. The Interaction between Allelic Variants of *CD86* and *CD40LG*: A Common Risk Factor of Allergic Asthma and Rheumatoid Arthritis. Allergy Asthma Immunol Res. 2014; 6:137−41.
177. Lenartowski R, Goc A. Tissue-specific association of the human tyrosine hydroxylase gene with the nuclear matrix. Neurosci Lettr. 2002; 330:151–4.
178. Leung T, Ko F, Sy H, Tsui S, Wong G. Differences in asthma genetics between Chinese and other populations. J Allergy Clin Immunol. In practice. 2014; 133:42−48.
179. Lewis C. Genetic association studies: design, analysis and interpretation. Briefings in Bioinformatics. 2002; 3:146−53.
180. Li G, Tolstonog G, Traub P. Interaction in vitro of type III intermediate filament proteins with Z-DNA and B-Z-DNA junctions. DNA Cell Biol. 2003; 22:141–69.

181. Li XJ, Li S. Proteasomal dysfunction in aging and Huntington disease. Neurobiol Dis. 2011; 43(1):4–8.
182. Li T, Kung H, Mack P, Gandara D. Genotyping and genomic profiling of non-small-cell lung cancer: implications for current and future therapies. Journal of Clinical Oncology. 2013; 31:1039–49.
183. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25:1451−2.
184. Liu P, Vikis H, Wang D, Lu Y, Wang Y, Schwartz A, et al. Familial aggregation of common sequence variants on 15q24-25.1 in lung cancer. Journal of the National Cancer Institute 2008; 100(18):1326–30.
185. Liu Y, Liu P, Wen W, James MA, Wang Y, Bailey–Wilson JE, et al. Haplotype and cell proliferation analyses of candidate lung cancer susceptibility genes on chromosome 15q24-25.1. Cancer research 2009; 69(19):7844–50.
186. Liu X, Wang X, Shen Y, Wu L, Ruan X, Lindpaintner K, et al. The functional variant rs1048990 in *PSMA6* is associated with susceptibility to myocardial infarction in a Chinese population. Atherosclerosis. 2009; 206:199–203.
187. Liu J, Yuan X, Liu J, Tian L, Quan J, Liu J, Chen X, et al.Validation of the association between PSMA6 -8 C/G polymorphism and type 2 diabetes mellitus in Chinese Dongxiang and Han populations. Diabetes Res Clin Pract. 2012; 98:295–301
188. Loisel D, Tan Z, Tisler C, Evans M, Gangnon R, Jackson D, et al. *IFNG* genotype and sex interact to influence the risk of childhood asthmaJ Allergy Clin Immunol. In practice. 2011; 128: 524−31.
189. Loos RJ. Recent progress in the genetics of common obesity. Br J Clin Pharmacol. 2009; 68(6):811–29.
190. Lou H, Yeager M, Li H, Bosquet J, Hayes R, Orr N, et al. Fine mapping and functional analysis of a common variant in *MSMB* on chromosome 10q11.2 associated with prostate cancer susceptibility. Proc Natl Acad Sci. 2009;106:7933–38.
191. Lv Y, Yan B, Yang H. Liu J, Zhong W, Liet K al. *LMP2/LMP7* gene variant: a risk factor for intestinal Mycobacterium tuberculosis infection in the Chinese population. Journal of Gastroenterology and Hepatology. 2011; 26(7):1145–50.
192. Mabb A, Judson M, Zylka M, Philpot B. Angelman syndrome: insights into genomic imprinting and neurodevelopmental phenotypes. Trends Neurosci. 2011; 34(6):293–303.
193. Maier L, Lowe C, Cooper J, Downes K, Anderson D, et al. IL2RA genetic heterogeneity in multiple sclerosis and type 1 diabetes susceptibility and soluble interleukin-2 receptor production. PLoS Genet. 2009; 5:e1000322.
194. Majetschak M, Sorell L, Patricelli T, Seitz D, Knoferl M. Detection and possible role of proteasomes in the bronchoalveolar space of the injured lung. Physiological research Academia Scientiarum Bohemoslovaca. 2008; 58:363−72.
195. Male V, Nisoli I, Gascoyne D, Brady H. *E4BP4:* an unexpected player in the immuneresponse.Trends Immunol 2012; 33:98e102.
196. Malerba G, Patuzzo C, Trabetti E, Lauciello M, Galavotti R, Pescollderungg L, et al. Chromosome 14 linkage analysis and mutation study of 2 serpin genes in allergic asthmatic families. J Allergy Clin Immunol. In practice. 2001; 107:654−58.

197. Mandel S, Fishman-Jacob T, Youdim M. Targeting SKP1, an ubiquitin E3 ligase component found decreased in sporadic Parkinson's disease. Neurodegener Dis. 2012; 10(1-4):220–223
198. Manly K. Reliability of statistical associations between genes and disease. Immunogenetics 2005; 57:549−58.
199. Mansur A, Bishop D, Holgate S, Markham A, Morrison J. Linkage/association study of a locus modulating total serum IgE on chromosome 14q13–24 in families with asthma. Thorax. 2004; 59:876−82.
200. Mayo I, Arribas J, Villoslada P, DoForno R-A, Rodriguez-Vilarifio S, et al. The proteasome is a major autoantigen in multiple sclerosis. Brain. 2002; 125:2658–67.
201. Meng L, Mohan R, Kwok B, Elofsson M, Sin N, Crews C. Epoxomicin, a potent and selective proteasome inhibitor, exhibits *in vivo* antiinflammatory activity. Proc Natl Acad Sci U S A. 1999; 96:10403-08.
202. Micale L, Fusco C, Augello B, Napolitano LM, Dermitzakis ET, Meroni G, et al. Williams-Beuren syndrome TRIM50 encodes an E3 ubiquitin ligase. Eur J Hum Genet. 2008; 16(9):1038–49.
203. Michlewski G, Guil S, Ca´ceres JF. Stimulation of pri-miR-18a processing by hnRNP A1. Adv Exp Med Biol. 2010; 700:28e35.
204. Mihailova A, Mikazane H, Klovins J, Nikitina-Zake L Association of protein tyrosine phosphatase non-receptor 22 (*PTPN22*) rs2476601 and Kruppel-like factor 12 (*KLF12*) rs1324913 single nucleotide polymorphisms with rheumatoid arthritis in a Latvian population. Scand J Rheumatol. 2011; 40:491–92.
205. Miller A. Multiple sclerosis: Where will be in 2020? Mt.Sinai J Med. 2011; 78:268–79.
206. Millward M, Price T, Townsend A, Sweeney C, Spencer A, Sukumaran S, et al. Phase 1 clinical trial of the novel proteasome inhibitor marizomib with the histone deacetylase inhibitor vorinostat in patients with melanoma, pancreatic and lung cancer based on in vitro assessments of the combination. Investigational new drugs. 2012; 30(6):2303–17.
207. Miltenberger-Miltenyi G, Laccone F. Mutations and polymorphisms in the human methyl CpG-binding protein MECP2. Hum. Mutat. 2003; 22(2) 107–15.
208. Minagar A, Ma W, Zhang X, Wang X, Zhang K, et al. Plasma ubiquitin-proteasome system profile in patients with multiple sclerosis: correlation with clinical features, neuroimaging, and treatment with interferon-beta-1b. Neurol Res. 2012; 34: 611–18.
209. Mishto M, Bellavista E, Ligorio C, Textoris-Taube K, Santoro A, et al. Immunoproteasome *LMP2* 60HH variant alters MBP epitope generation and reduces the risk to develop multiple sclerosis in Italian female population. PLoS One. 2010; 5:e9287.
210. Mohamed R, Pasha H, El-Shahawy E Influence of *TRAF1/C5* and *STAT4* genes polymorphisms on susceptibility and severity of rheumatoid arthritis in Egyptian population. Cell Immunol. 2012; 273:67–72.
211. Morgan A, Robinson J, Conaghan P, Martin S, Hensor E, et al. Evaluation of the rheumatoid arthritis susceptibility loci *HLA-DRB1, PTPN22, OLIG3/TNFAIP3*, STAT4 and *TRAF1/C5* in an inception cohort. Arthritis Res Ther. 2010; 12:R57.
212. Mountz J. Significance of increased circulating proteasome in autoimmune disease. J Rheumatol. 2002; 29:2027–30.

213. Moutzouris JP, Che W, Ramsay EE, Manetsch M, Alkhouri H, Bjorkman AM, et al. Proteasomal inhibition upregulates the endogenous *MAPK* deactivator *MKP-1* in human airway smooth muscle: mechanism of action and effect on cytokine secretion. Biochimica et biophysica acta 2010;1803(3):416–23.
214. Munthe-Kaas MC, Gerritsen J, Carlsen KH, Undlien D, Egeland T, Skinningsrud B, Tørres T, Carlsen KL. Eosinophil cationic protein (ECP) polymorphisms and association with asthma, s-ECP levels and related phenotypes. Allergy. 2007; 62(4):429–36.
215. Nagy G, Koncz A, Telarico T, Fernandez D, Ersek B, et al. Central role of nitric oxide in the pathogenesis of rheumatoid arthritis and systemic lupus erythematosus. Arthritis Res Ther. 2010; 12:210.
216. Nalepa G, Rolfe M, Harper J. Drug discovery in the ubiquitin-proteasome system. Nat Rev Drug Discov. 2006; 5:596–13.
217. Nandi D, Tahiliani P, Kumar A, Chandu D. The ubiquitin-proteasome system. Journal of biosciences 2006; 31(1):137–55.
218. Naujokat C, Hoffmann S. Role and function of the 26S proteasome in proliferation and apoptosis. Laboratory investigation; a journal of technical methods and pathology 2002; 82(8):965–80.
219. Nawrocki S, Carew J, Dunner K Jr, et al. Bortezomib inhibits PKR-like endoplasmic reticulum (ER) kinase and induces apoptosis *via* ER stress in human pancreatic cancer cells. Cancer Res. 2005; 65:11510−19.
220. Ng P, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res. 2001; 11:863–74.
221. Nicolae D, Gamazon E, Zhang W, Duan S, Dolan M, Cox N. Traitassociated SNPs are more likely to be eQTLs: Annotation to enhance discovery from GWAS. PLoS Genet. 2010; 6:e1000888.
222. Nijhawan D, Zack T, Ren Y et al. Cancer vulnerabilities unveiled by genomic loss. Cell. 2012; 150:842–54.
223. Nokoff N, Rewers M. Pathogenesis of type 1 diabetes: lessons from natural history studies of high-risk individuals. Ann N Y Acad Sci. 2013; 1281:1–15.
224. Olsen N, Stein C. New drugs for rheumatoid arthritis. N Engl J Med. 2004; 350:2167–79.
225. O'Rielly D, Roslin N, Beyene J, Pope A, Rahman P. *TNF*-alpha-308 G/A polymorphism and responsiveness to *TNF*-alpha blockade therapy in moderate to severe rheumatoid arthritis: a systematic review and meta-analysis. Pharmacogenomics J. 2009; 9:161–67.
226. Orlowski M. The multicatalytic proteinase complex, a major extralysosomal proteolytic system. Biochemistry. 1990; 29:10289–97.
227. Otoda T, Takamura T, Misu H, Ota T, Murata S, et al. Proteasome dysfunction mediates obesity-induced endoplasmic reticulum stress and insulin resistance in the liver. Diabetes. 2013; 62:811–24.
228. Owen S, Lunt M, Bowes J, Hider SL, Bruce I, et al. *MTHFR* gene polymorphisms and outcome of methotrexate treatment in patients with rheumatoid arthritis: analysis of key polymorphisms and meta-analysis of *C677T* and *A1298C* polymorphisms. Pharmacogenomics J. 2013; 13:137–47.
229. Ozaki K, Sato H, Iida A, Mizuno H, Nakamura T, Miyamoto Y, et al. A functional SNP in *PSMA6* confers risk of myocardial infarction in the Japanese population. Nature Genet. 2006; 38:921−25.

230. Padyukov L, Seielstad M, Ong R, Ding B, Rönnelid J, et al. A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis. Ann Rheum Dis. 2011; 70:259–65.
231. Palombella V, Rando O, Goldberg A, Maniatis T. The ubiquitin-proteasome pathway is required for processing the NF-κB1 precursor protein and the activation of NF-κB. Cell. 1994;78:773−85.
232. Pan H, Chang D, Feng L, Xu F, Kuang H, et al. Oxidative damage to DNA and its relationship with diabetic complications. Biomed Environ Sci. 2007; 20:160–63.
233. Papandreou C, Daliani D, Nix D, et al. Phase I trial of the proteasome inhibitor bortezomib in patients with advanced solid tumors with observations in androgen-independent prostate cancer. J Clin Oncol. 2004;22:2108−21.
234. Paul D, Nisbet J, Yang T, Meacham S, Rendon A, Hautaviita K, et al. Maps of open chromatin guide the functional follow-up of genome-wide association signals: Application to hematological traits. PLoS Genet. 2011; 7:e1002139.
235. Paz Z, Tsokos G. New therapeutics in systemic lupus erythematosus. Current Opinion in Rheumatology. 2013; 25:297–303
236. Petty R, Southwood T, Manners P, Baum J, Glass D, Goldenberg J, et al. International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, J Rheumatol 2004; 31:390−2.
237. P´erez V, Buffenstein R, Masamsetti V, Leonard S, SalmonA, MelebJ, et al. Protein stability and resistance to oxidative stress are determinants of longevity in the longest-living rodent, the naked mole-rat. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106(9):3059–64.
238. Phillips K, Luisi B. The virtuoso of versatility: POU proteins that flex to fit. J Mol Biol 2000; 302:1023e39.
239. Pique-Regi R, Degner J, Pai A, Gaffney D, Gilad Y, Pritchard J. Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. Genome Res. 2011; 21:447–55.
240. Pitzalis M, Zavattari P, Murru R, Deidda E, Zoledziewska M, Murru D, et al. Genetic loci linked to type 1 diabetes and multiple sclerosis families in Sardinia. BMC Med Genet. 2008; 9:3.
241. Plant D, Bowes J, Potter C, Hyrich K, Morgan A, et al. Genome–wide association study of genetic predictors of anti-tumor necrosis factor treatment efficacy in rheumatoid arthritis identifies associations with polymorphisms at seven loci. Arthritis Rheum. 2011; 63:645–53.
242. Prahalad S, Glass DN. A comprehensive review of the genetics of juvenile idiopathic arthritis. Pediatr Rheumatol Online J. 2008; 6:11.
243. Prasanth S, Chattopadhyay M, Bhat K, Ali S. Expression of protooncogene c-kit receptor in rat testis and uniqueness of extracellular domain across the species with potential in molecular phylogeny. DNA Cell Biol. 2004; 23(1):35–43.
244. Pryhuber K, Murray K, Donnelly P, Passo M, Maksymowych W, Glass D, et al. Polymorphism in the *LMP2* gene influences disease susceptibility and severity in *HLAB27* associated juvenile rheumatoid arthritis. J Rheumatol 1996; 23: 747–52.

245. Qian F, Zhang Q, Zhou L, Ma G, Jin G, Huang Q, et al. Association between polymorphisms in *IL17F* and male asthma in a Chinese population. J Investig Allergol Clin Immunol. 2012; 22(4):257−63.
246. Reich N. STATs get their move on. JAKSTAT. 2013; 2:e27080.
247. Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 2000; 132:365−86.
248. Richardson P, Mitsiades C, Hideshima T, Anderson K. Bortezomib: proteasome inhibition as an effective anticancer therapy. Annu Rev Med. 2006; 57:33-47.
249. Raby B, Lazarus R, Silverman E, Lake S, Lange C, Wjst M. Association of vitamin D receptor gene polymorphisms with childhood and adult asthma. Am J Respir Crit Care Med 2004; 170(10):1057−65.
250. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 2002; 30(17):3894–900.
251. Ramírez-Bello J, Jiménez-Morales S, Espinosa-Rosales F, Gómez-Vera J, Gutiérrez A, Velázquez Cruz R, et al. Juvenile rheumatoid arthritis and asthma, but not childhood-onset systemic lupus erythematosus are associated with *FCRL3* polymorphisms in Mexicans. Mol Immunol. 2013; 53(4):374–8.
252. Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2009;369:767−78.
253. Richardson P, Mitsiades C, Hideshima T, Anderson K. Bortezomib: proteasome inhibition as an effective anticancer therapy. Annu Rev Med. 2006; 57:33–47.
254. Ridolfi E, Fenoglio C, Cantoni C, Calvi A, De Riz M, et al. Expression and Genetic Analysis of MicroRNAs Involved in Multiple Sclerosis. Int J Mol Sci. 2013; 14:4375–84.
255. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, et al. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nat Methods. 2007; 4:651–57.
256. Rodenhiser D., Mann M. Epigenetics and human disease: translating basic biology into clinical application. CMAJ. 2006; 174(3):341–48.
257. Rodriguez S, Gaunt T, Day I. Hardy-Weinberg equilibrium testing of biological ascertainment for Mendelian randomization studies. Am. J. Epidemiol. 2009; 169:505–14.
258. Rojas-Villarraga A, Amaya-Amaya J, Rodriguez-Rodriguez A, Mantilla RD, Anaya JM. Introducing polyautoimmunity: secondary autoimmune diseases no longer exist. Autoimmune Dis 2012; 254319.
259. Rothenburg S, Koch-Nolte F, Haag F. DNA methylation and Z-DNA formation as mediators of quantitative differences in the expression of alleles. Immunol Rev. 2001;184:286–98.
260. Rozas J, Sánchez - DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 2003; 19(18): 2496-97.
261. Saccone S, Bolze R, Thomas P, Quan J, Mehta G, Deelman E, et al. SPOT: A web-based tool for using biological databases to prioritize SNPs after a genome-wide association study. Nucleic Acids Res. 2010; 38:W201–W209.
262. Sánchez-Zamora YI, Rodriguez-Sosa M. The role of MIF in type 1 and type 2 diabetes mellitus. J Diabetes Res. 2014;804519.
263. Sandholm N, Salem R, McKnight A, Brennan E, Forsblom C et al. New susceptibility loci associated with kidney disease in type 1 diabetes. PLoS Genet. 2012; 8:e1002921.

264. Sakamoto K, Sato Y, Sei M, Ewis A, Nakahori Y. Proteasome activity correlates with male BMI and contributes to the differentiation of adipocyte in hADSC. Endocrine. 2010; 37:274–79.
265. Schadt E, Molony C, Chudin E, Hao K, Yang X, Lum P, Kasarskis et al. Mapping the genetic architecture of gene expression in human liver. PLoS Biol 2008; 6:e107.
266. Scheneider-Hohendorf T, Stenner M, Weidenfeller C, Zozulya A. Regulatory T cells exhibit enhanced migratory characteristics, a feature impaired in patients with multiple sclerosis. Eur J Immunol. 2010; 40:3581–90.
267. Schmidt H. Introducing single-nucleotide polymorphism markers in the diagnosis of Wilson disease. Clinical Chemistry. 2007; 53(9):1568–69.
268. Schork N, Fallin D, Lanchbury J. Single nucleotide polymorphisms and the future of genetic epidemiology. Clin. Genet. 2000; 58:250–64.
269. Schwartz A, Ciechanover A. Targeting proteins for destruction by the ubiquitin system: implications for human pathobiology. Annual review of pharmacology and toxicology 2009; 49:73–96.
270. Scibetta A, Santangelo S, Coleman J, Hall D, Chaplin T, Copier J, et al. Functional analysis of the transcription repressor PLU-1/JARID1B. Mol Cell Biol. 2007; 27:7220–35.
271. Seibold M, Wang B, Eng C, Kumar G, Beckman K, Sen S, et al. An African-specific functional polymorphism in *KCNMB1* shows sex–specific association with asthma severity. Hum Mol Gen. 2008; 17(17):2681−90.
272. Selvaraju V, Joshi M, Suresh S, Sanchez JA, Maulik N, et al. Diabetes, oxidative stress, molecular mechanism, and cardiovascular disease- an overview. Toxicol Mech Methods. 2012; 22:330-35.
273. Shah I, Lees K, Pien C, Elliott P. Early clinical experience with the novel proteasome inhibitor PS-519. British journal of clinical pharmacology 2002;n54(3):269–76.
274. Sharma A, Bernatchez P, de Haan J. Targeting endothelial dysfunction in vascular complications associated with diabetes. Int J Vasc Med. 2012; 750126.
275. Sharp L, Little J. Polymorphisms in genes involved in folate metabolism and colorectal neoplasia: A HuGE Review. Am. J. Epidemiol. 2004: 159:423–43.
276. She H, Mao Z. Regulation of myocyte enhancer factor-2 transcription factors by neurotoxins. Neurotoxicology 2011; 32:563e6.
277. Sivakumaran S, Agakov F, Theodoratou E, Prendergast J, Zgaga L, Manolio T, et al. Abundant pleiotropy in human complex diseases and traits. Am J Hum Genet. 2011; 89(5):607−18.
278. Sixt S, Peters J. Extracellular alveolar proteasome: possible role in lung injury and repair. Proceedings of the American Thoracic Society. 2010; 7:91−96.
279. Sjakste T, Sjakste N, Scherrer K. Exon/intron organisation of human proteasome *PROS–27 K* gene. DNA Seq. 2001; 12: 261–5.
280. Sjakste T, Lauberte L, Collan Y, Savontaus M, Bajare A, et al. Identification of an intronic TG repeat polymorphism in the human proteasome core particle *PROS-27K* gene. DNA Seq. 2002; 13:139–43.
281. Sjakste T, Eglite J, Sochnevs A, Marga M, Pirags V, et al. Microsatellite genotyping of chromosome 14q13.2-14q13 in the vicinity of proteasomal gene *PSMA6* and association with Graves' disease in the Latvian population. Immunogenetics. 2004a; 56238–43.
282. Sjakste T, Kalis M, Poudziunas I, Pirags V, Lazdins M, et al. Association of microsatellite

polymorphisms of the human 14q13.2 region with type 2 diabetes mellitus in Latvian and Finnish populations. Ann Human Genet. 2004b; 71:772–76.
283. Sjakste T, Poudziunas I, Ninio E, Perret C, Pirags V, et al. SNPs of *PSMA6* gene - investigation of possible association with myocardial infarction and type 2 diabetes mellitus. Russ J Genet. 2007; 43:442–48.
284. Sjakste T, Poudziunas I, Pirags V, Lazdins M, N. Sjakste Bioinformatical analysis of evolutional conservatism and functional significance of microsatellite alleles of human 14q13.2 region associated with type 2 diabetes mellitus. Proc. Latv. Acad. Sci. B 2008; 62:91–102.
285. Sjakste T, Trapina I, Rumba-Rozenfelde I, Lunin R, Sugoka O, et al. 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. 2010; 29:543–51.
286. Själander A, Birgander R, Saha N, Beckman L, Beckman G. p53 polymorphisms and haplotypes show distinct differences between major ethnic groups. Hum Hered. 1996; 46(1):41–8.
287. Sjöblom T, Jones S, Wood L, Parsons D, Lin J, Barber T, et al. The consensus coding sequences of human breast and colorectal cancers. Science. 2006; 314(5797):268–74.
288. Smith KJ, Lassmann H The role of nitric oxide in multiple sclerosis. Lancet Neurol. 2002; 1:232–41.
289. Soejima M, Kawaguchi Y, et al. Prospective study of the association between *NAT2* gene haplotypes and severe adverse events with sulfasalazine therapy in patients with rheumatoid arthritis. J Rheumatol. 2008; 35:724.
290. Song G, Choi S, Ji J, Lee Y. Genome-wide pathway analysis of a genome-wide association study on multiple sclerosis. Mol Biol Rep. 2013; 40:2557–64.
291. Soon W, Hariharan M, Snyder M. High-throughput sequencing for biology and medicine. Molecular Systems Biology. 2013; 9:640.
292. Sorokin A, Kim E, Ovchinnikov L. Proteasome system of protein degradation and processing. Biochemistry (Moskow). 2009; 74: 1411–42.
293. Stahl E, Raychaudhuri S, Remmers E, Xie G, Eyre S, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet. 2010;42:508–14.
294. Stranger B, Nica A, Forrest M, Dimas A, Bird C, Beazley C, et al. Population genomics of human gene expression. Nat Genet. 2007; 39:1217–24.
295. Stuart P, Nair R, Ellinghaus E, Ding J, Tejasvi T, Gudjonsson J, et al. Genome-wide association analysis identifies three psoriasis susceptibility loci. Nat Genet 2010; 42:1000e4.
296. Sun J, Helgason A, Masson G, Ebenesersdóttir S, Li H, Mallick S, et al. A direct characterization of human mutation based on microsatellites. Nat Genet. 2012; 44:1161–5.
297. Sutoh Y, Kondo M, Ohta Y, Ota T, Tomaru U, Flajnik M, et al. Comparative genomic analysis of the proteasome β5t subunit gene: implications for the origin and evolution of thymoproteasomes. Immunogenetics. 2012; 64:49–58.
298. Sutovsky P. Sperm proteasome and fertilization. Reproduction. 2011; 142(1):1–14.
299. Stys P, Zamponi G, va Minnen J, Geurts J Will the real multiple sclerosis stand up? Nat Rev Immunol. 2012; 13:507–14.

300. Tajima, F. Statistical method for testing the neutral mutationhypothesis by DNA polymorphism. Genetics. 1989; 123:585−95.
301. Takashima, N., Shioji, K., Kokubo, Y. Validation of the association between the gene encoding proteasome subunit α type 6 and myocardial infarction in a Japanese population. Circ. J. 2007; 71: 495–98.
302. Takatori RK, Takahashi A, Tokunaga D, Hojo T, Fujioka M, et al. *ABCB1* C3435T polymorphism influences methotrexate sensitivity in rheumatoid arthritis patients. Clin Exp Rheumatol. 2006; 24:546–54.
303. Tatsch E, Bochi G, Piva S, De Carvalho J, Kober H, et al. Associationbetween DNA strand breakage and oxidative, inflammatory and endothelial biomarkers in type 2 diabetes. Mutat Res. 2012; 732:16–20.
304. Tautz D, Schlötterer C. Simple sequences. Curr Opin Genet Dev. 1994; 4:832–7.
305. Thompson SD, Sudman M, Ramos PS. The susceptibility loci Juvenile Idiopathic Arthritis shares with other autoimmune diseases extend to *PTPN2, COG6* and *ANGPT1*. Arthritis Rheum. 2010; 62(11):3265–76.
306. Tolstonog G, Li G, Shoeman R, Traub P. Interaction in vitro of type III intermediate filament proteins with higher order structures of single-stranded DNA, particularly with G-quadruplex DNA. DNA Cell Biol. 2005; 24:85–110.
307. Tomaru U, Ishizu A, Murata S, Miyatake Y, Suzuki S, Takahashi S, et al. Exclusive expression of proteasome subunit {beta}5t in the human thymic cortex. Blood. 2009; 113:5186–91.
308. Tomer Y, Davies T. The genetic susceptibility to Graves' disease. Baillière's Clinical Endocrinology and Metabolism. 1997; 11(3):431–50
309. Tomko R, Jr. and M. Hochstrasser. Molecular architecture and assembly of the eukaryotic proteasome. Annual Review of Biochemistry. 2013; 82:415–45.
310. Torrelo A, Patel S, Colmenero I. Chronic atypical neutrophilic dermatosis with lipodystrophy and elevated temperature (CANDLE) syndrome. Journal of the American Academy of Dermatology. 2010; 62(3):489–95.
311. Trapina I, Rumba-Rozenfelde I, Sjakste N, Sokolovska J, Sugoka O et al. 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 2009; 63(4/5):20–30.
312. Tsukamoto S, Yokosawa H. Natural products inhibiting the ubiquitin-proteasome proteolytic pathway, a target for drug development. Curr Med Chem. 2006;13:745–54.
313. Tsygankov AY. TULA-family proteins: an odd couple. Cell Mol Life Sci. 2009; 66:2949–52.
314. Tulah A, Beghé B, Barton S, Holloway J, Sayers I. Leukotriene B4 receptor locus gene characterisation and association studies in asthma. BMC Med Genet. 2012; 13:110.
315. Van Endert P, Liblau R, Patel S, Fugger L, Lopez T, Pociot F, et al. Major histocompatibility complex-encoded antigen processing gene polymorphism in *IDDM*. Diabetes. 1994; 43:110–17.
316. Venken K, Hellings N, Hensen K, Rummens J, Stinissen P Memory CD4+CD127 high T cells from patients with multiple sclerosis produce IL-17 in response to myelin antigens. J Neuroimmunol. 2010; 226:185–91.

317. Vernace V, Schmidt-Glenewinkel T, Figueiredo-Pereira M. Aging and regulated protein degradation: who has the UPPer hand? Aging Cell. 2007; 6(5):599–606.
318. Voigt A, Bartel K, Vliet J, Oates N, Whitelaw E. Epigenetic mechanisms in context in the complex diseases. Cell. Mol. Life Sci.. 2007; 64:1531–38.
319. Voutsadakis IA. Ubiquitin, ubiquitination and the ubiquitinproteasome system in cancer. Atlas Genet Cytogenet Oncol Haematol. 2010; 14(11):1088–99.
320. Udvardy, A. Purification and characterization of a multiprotein component of the Drosophila 26 S (1500 kDa) proteolytic complex. J. Biol. Chem. 1993; 268:9055–62.
321. Ungvári I, Hullám G, Antal P, Kiszel P, Gézsi A, Hadadi É, et al. Evaluation of a partial genome screening of two asthma susceptibility reģions using bayesian network based bayesian multilevel analysis of relevance. PLoS. 2012; One 7:e33573.
322. Uthra S, Raman R, Mukesh B, Rajkumar S, Kumari P, et al. Diabetic retinopathy: Validation study of *ALR2, RAGE*, *iNOS* and *TNFB* gene variants in a south Indian cohort. Ophthalmic Genet. 2010; 31: 244–51.
323. Uzumcu A, Karaman B, Toksoy G, Uyguner ZO, Candan S, Eris H, et al. 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. 2009; 52:315e20.
324. Wang J, Maldonado MA. The ubiquitin-proteasome system and its role in inflammatory and autoimmune diseases. Cellular & molecular immunology 2006; 3(4):255–61.
325. Wang G, Vasquez K. Z-DNA, an active element in the genome. Front Biosci. 2007; 12:4424–38.
326. Wang X, Guerrero C, Kaiser P, Huang L. Proteomics of proteasome complexes and ubiquitinated proteins. Expert review of proteomics. 2007; 4:649−65.
327. Wang L, Kumar S, Fridley BL, Kalari K, Moon I, Pelleymounter L, et al. Proteasome beta subunit pharmacogenomics: gene resequencing and functional genomics. Clin. Cancer Res. 2008; 14:3503–13.
328. Wang H, Jiang M, Zhu H, Chen Q, Gong P, Lin J, et al. Quantitative assessment of the influence of *PSMA6* variant (rs1048990) on coronary artery disease risk. Mol Biol Rep. 2013; 40:1035−41.
329. Weber JL, Wong C. Mutation of human short tandem repeats. Hum Mol Genet 1993; 2:1123–8.
330. Williams C, Driscoll D, Dagli AI. Clinical and genetic aspects of Angelman syndrome. Genet Med. 2010; 12(7):385–95.
331. Wilson S, Bhattacharyya B, Rachel R, Coppola V, Tessarollo L, Householder D, et al. Synaptic defects in ataxia mice result from a mutation in Usp14, encoding a ubiquitin-specific protease. Nat Genet. 2002; 32(3):420–25.
332. Wood A,. Perry J, Tanaka T. Hernandez D, Zheng HF, et al. Imputation of variants from the 1000 Genomes Project modestly improves known associations and can identify low–frequency variant-phenotype associations undetected by HapMap based imputation. PLoS ONE. 2013; 8:e64343.
333. Wu L, Sjakste T, Sakalauskas R, Sitkauskiene B, Paramonova N, Gasiuniene E, et al. The burden of allergic asthma in children: a landscape comparison based on data from Lithuanian, Latvian, and Taiwanese populations. Pediatrics and neonatology. 2012; 53:276−82.

334. Wu Z, Shi Y. When ubiquitin meets NF-κB: a trove for anti-cancer drug development. Curr Pharm Des. 2013; 19(18):3263–75.
335. Xu H, Lu A, Sharp F. BMC Genomics. 2011; 12:499.
336. Xu Z, Taylor J. SNPinfo: Integrating GWAS and candidate geneinformation into functional SNP selection for genetic association studies. Nucleic Acids Res. 2009; 37:W600–W605.
337. Yang H, Chang L, Liang Y, Lin C, Wang P. A genome-wide homozygosity association study identifies runs of homozygosity associated with rheumatoid arthritis in the human major histocompatibility complex. PLoS One. 2012; 7(4):e34840.
338. Yang I, Schwartz D. Epigenetic mechanisms and the development of asthma. J. Allergy Clin. Immunol. 2012; 130(6):1243–55.
339. Zemeckienė Z, Vitkauskienė A, Sjakste T, Sitkauskienė B, Sakalauskas R. Proteasomes and proteasomal gene polymorphism in association with inflammation and various diseases. Medicina (Kaunas). 2013; 49:207–13.
340. Zhang Y, Moffatt M, Cookson W. Genetic and genomic approaches to asthma: new insights for the origins. Curr Opin Pulm Med. 2012; 18(2):6−13.
341. Zhang K, Chang S, Cui S, Guo L, Zhang L, et al. ICSN Pathway: identify candidate causal SNPs and pathways from genome-wide association study by one analytical framework. Nucleic Acids Res. 2011; 39:W437–43.
342. Zheng J, Bizzozero O. Decreased activity of the 20S proteasome in the brain white matter and gray matter of patients with multiple sclerosis. Journal of Neurochemistry. 2011; 117:143–53.
343. Zhong H, Beaulaurier J, Lum P, Molony C, Yang X, Macneil D,Weingarth DT, Zhang B, Greenawalt D, Dobrin R, et al. Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. PLoS Genet. 2010; 6:e1000932.
344. Zoledziewska M, Costa G, Pitzalis M, Cocco E, Melis C, et al. Variation within the CLEC16A gene shows consistent disease association with both multiple sclerosis and type 1 diabetes in Sardinia. Genes Immun. 2009; 10:15–17.

ACKNOWLEDGEMENTS

I would like to thanks, first of all, my supervisor Prof. Nikolajs Sjakste for supervising the work. Special thanks to Dr. Biol. Tatjana Sjakste for possibility to perform this work in her research "Genomics and bioinformatics" group of Latvian University of agency "LU Institute of Biology," for her helpful advises and consultation.

Also like to thank the staff of research group "Genomics and bioinformatics" in the Institute of Biology of the University of Latvia and Biochemistry lecturer group (Latvian University, Faculty of Medicine) team for their assistance and support work.

I would and a big thank i also wants to express my parents for support through education, not too light path.

In conclusion would also like to thank Dr. biol. Maja Dzintare of amber for text editing Latvian.

The work was carried out with the ESF project Nr.2013 / 0043 / 1DP / 1.1.1.2.0 / 13 / APIA / SEDA / 002 support.