doi:10.25143/prom-rsu_2013-12_dt

RĪGAS STRADIŅA UNIVERSITĀTE

Inga Kempa

IDENTIFICATION OF CANDIDATE GENES INVOLVED IN THE ETIOLOGY OF NON-SYNDROMIC CLEFT LIP WITH OR WITHOUT CLEFT PALATE AND ISOLATED CLEFT PALATE

Doctoral Thesis
for obtaining the degree of a Doctor of Medicine
Speciality – Medical Genetics

Rīga, 2013

RĪGAS STRADIŅA UNIVERSITĀTE

Inga Kempa

IDENTIFICATION OF CANDIDATE GENES INVOLVED IN THE ETIOLOGY OF NON-SYNDROMIC CLEFT LIP WITH OR WITHOUT CLEFT PALATE AND ISOLATED CLEFT PALATE

Doctoral Thesis

for obtaining the degree of a Doctor of Medicine

Speciality – Medical Genetics

Scientific supervisors:

Dr. med. **Baiba Lāce**

Dr. biol. **Jānis Kloviņš**

Riga, 2013

ANNOTATION

Cleft lip with or without cleft palate and isolated cleft palate (CL/CLP/CP) is one of the most common birth defects worldwide with prevalence of approximately 1 in 700 live births in European populations. Individuals with CL/CLP/CP need multidisciplinary care from birth to adulthood even after surgical repair. CL/CLP/CP affects speech, dental development, hearing, appearance and psychology of person. It has been considered that individuals with this malformation have higher morbidity and mortality of cardiovascular diseases and cancers if compared to unaffected individuals. Despite possibility of surgical repair, this defect remains important health and social problem in nowadays.

Formation of orofacial clefts is a result between interaction of environmental and genetic factors. Recent estimates suggest that 2-14 genes could be involved in the formation of CL/CLP/CP. In present study we performed case-control analysis and family based association test in CL/CLP and CP patients, their parents and control group. Identification of possible candidate genes involved in the etiology of non-syndromic cleft lip with or without cleft palate and isolated cleft palate in Latvian population was the objective of the present study. Our results showed very strong association between *FGFR1*, *WNT3*, *SKI*, *BMP4* and *IRF6* genes and non-syndromic CL/CLP and CP and possible interaction between 19q13 locus and non-syndromic CL/CLP, which continue to support the involvement of these genes in the development of non-syndromic clefts in Caucasians.

Results of this study is step further of understanding of this complex malformation and estimating the impact of genes involved in the etiology of non-syndromic cleft lip with or without cleft palate and isolated cleft palate.

ANOTĀCIJA

Lūpas šķeltne ar/bez aukslēju šķeltnes un izolēta aukslēju šķeltne (LŠ/LŠ+AŠ/AŠ) ir viens no visbiežāk sastopamajiem iedzimtajiem defektiem visā pasaulē, ar prevalenci Eiropas populācijā aptuveni 1 no 700 jaundzimušajiem. Indivīdiem ar LŠ/LŠ+AŠ/AŠ nepieciešama multidisciplināra aprūpe visā dzīves laikā, arī pēc ķirurģiskas operācijas. Lūpas šķeltne ar/bez aukslēju šķeltnes un izolēta aukslēju šķeltne ir saistīta ne tikai ar personas runas, dzirdes traucējumiem, zobu attīstības, bet arī ar izskata un psiholoģiskām problēmām. Tiek uzskatīts, ka indivīdiem ar LŠ/LŠ+AŠ/AŠ, ir paaugstināts sirds-asinsvadu slimību un audzēju saslimstības un mirstības risks, salīdzinot ar veseliem indivīdiem. Neskatoties uz iespēju ķirurģiski labot šo defektu, mūsdienās LŠ/LŠ+AŠ/AŠ ir kļuvusi par vienu no svarīgām sabiedrības veselības problēmām visā pasaulē.

LŠ/LŠ+AŠ/AŠ veidojas, mijiedarbojoties ārējās vides faktoriem un ģenētiskajiem faktoriem. Pēdējie pētījumi liecina, ka aptuveni 2-14 gēni varētu būt iesaistīti nesindromālo LŠ/LŠ+AŠ/AŠ veidošanā. Šajā pētījumā tika veikta gadījuma-kontroles analīze un ģimenes asociācijas tests, lai identificētu iespējamos kandidātgēnus, kuri varētu būt iesaistīti nesindromālo lūpas šķeltnes ar/bez aukslēju šķeltnes un izolētas aukslēju šķeltnes attīstībā Latvijas populācijā.

Mūsu pētījumā iegūtie rezultāti atklāj ļoti augstu saistību starp *FGFR1*, *WNT3*, *SKI*, *BMP4* un *IRF6* gēniem un nesindromālajām LŠ/LŠ+AŠ un AŠ, kā arī norāda uz iespējamo saistību starp 19q13 lokusu un nesindromālajām LŠ/LŠ+AŠ. Šie rezultāti turpina apstiprināt minēto gēnu nozīmi nesindromālo lūpas ar/bez aukslēju šķeltņu un izolētas aukslēju šķeltnes attīstībā eiropiešiem.

Šis ir pirmais tik liela mēroga pētījums, kas ir veltīts nesindromālo LŠ/LŠ+AŠ/AŠ kandidātgēnu analīzei Latvijā. Pētījumā iegūtie rezultāti ir vērā ņemams ieguldījums šīs sarežģītās patoloģijas izpratnē un iespējamo gēnu ietekmes izvērtēšanā slimības izraisīšanā.

TABLE OF CONTENTS

ABBREVIATION

ADH1C	-	alcohol dehydrogenase 1C (class I), gamma polypeptide
APEX-2	-	Arrayed primer extension reaction-2
APOC2	-	apolipoprotein C-I
BCL3	-	B-cell CLL/lymphoma 3
BMP2	-	bone morphogenetic protein 2
BMP4	-	bone morphogenetic protein 4
camp	-	cathelicidin antimicrobial peptide
CDH1	-	cadherin 1, type 1, E-cadherin (epithelial)
CEU	-	U.S. residents (Utah) with northern and western European ancestry
CI	-	confidence interval
CL	-	cleft lip
CL/CLP	-	cleft lip and cleft lip with cleft palate
CL/CLP/CP	-	cleft lip with or without cleft palate
CP	-	cleft palate
CLP	-	cleft lip with cleft palate
CLPTM1	-	cleft lip and palate associated transmembrane protein 1
COL11A1	-	collagen, type XI, alpha 1
COL11A2	-	collagen, type XI, alpha 2
COL2A1	-	collagen, type II, alpha 1
COL9A1	-	collagen, type IX, alpha 1
COL9A2	-	collagen, type IX, alpha 2
DMSO	-	Dimethyl sulfoxide
DNA	-	Deoxyribonucleic acid
dNTP	-	Deoxynucleotide Triphosphate
DZ	-	dizygotic twins
EDN1	-	endothelin 1
EDTA	-	Ethylenediaminetetraacetic acid
EMT	-	epithelial-mesenchymal transformation
FGF1	-	fibroblast growth factor 1 (acidic)
FGF10	-	fibroblast growth factor 10
FGF16	-	fibroblast growth factor 16

FGF2 - fibroblast growth factor 2

FGF23 - fibroblast growth factor 23

FGFR1 - fibroblast growth factor receptor 1

FGFR4 - fibroblast growth factor receptor 4

FN1 - fibronectin 1

FOXE1 - forkhead box E1 (thyroid transcription factor 2)

GLI2 - GLI family zinc finger 2

GSTT1 - glutathione S-transferase theta 1

HWE - Hardy Weinberg equilibrium

IHH - isolated hypogonadotropic hypogonadism

IRF6 - interferon regulatory factor 6

JAG2 - jagged 2

LHX8 - LIM homeobox 8

MAF - minor allele frequency

MALDI-TOF - Matrix-assisted laser desorption/ionization-time-of-flight

MFT - multifactorial treshold model

$MgCl_2$ - magnesium chloride

MMP13 - matrix metallopeptidase 13 (collagenase 3)

MMP2 - matrix metallopeptidase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)

MMP25 - matrix metallopeptidase 25

MMP3 - matrix metallopeptidase 3 (stromelysin 1, progelatinase)

MMP9 - matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)

MSX1 - msh homeobox 1

MSX2 - msh homeobox 2

MTHFR - methylenetetrahydrofolate reductase (NAD(P)H)

MZ - monozygotic twins

NaCl - sodium chloride

NaOH - sodium hydroxide

NE buffer - elution buffer

NOS3 - nitric oxide synthase 3 (endothelial cell)

NT1 buffer - binding buffer

NT3 buffer - washing buffer
NUDT6 - nudix (nucleoside diphosphate linked moiety X)-type motif 6
OR - odds ratio
OSMED - otospondylomegaepiphyseal dysplasia
PCR - polymerase chain reaction
PVR - poliovirus receptor
PVRL1 - poliovirus receptor-related 1 (herpesvirus entry mediator C)
PVRL2 - poliovirus receptor-related 2 (herpesvirus entry mediator B)
RARA - retinoic acid receptor, alpha
RBC Lysis A1 solution
SAP - Shrimp Alkaline Phosphatase
SATB2 - SATB homeobox 2
SDS - Sodium dodecyl sulfate
SKI - v-ski sarcoma viral oncogene homolog (avian)
SLS-1 - saliva lysis solution
SMAD2 - SMAD family member 2
SMAD3 - SMAD family member 3
SMAD4 - SMAD family member 4
SNP - single nucleotide polymorphism
SPRY2 - sprouty homolog 2 (Drosophila)
SYN3 - synapsin III
TBX10 - T-box 10
TBX22 - T-box 22
TDT - transmission disequilibrium test
TE buffer - Tris-EDTA buffer
TGFA - transforming growth factor, alpha
TGFB3 - transforming growth factor, beta 3
TIMP1 - TIMP metallopeptidase inhibitor 1
TIMP2 - TIMP metallopeptidase inhibitor 2
TIMP3 - TIMP metallopeptidase inhibitor 3
Tris-HCl - tris(hydroxymethyl)aminomethane
WB1 buffer - washing buffer 1
WB2 buffer - washing buffer 2

EB buffer	- elution buffer
WHO	- world health organization
WNT3	- wingless-type MMTV integration site family, member 3
WNT9B	- wingless-type MMTV integration site family, member 9B
WZS	- Weissenbach-Zweymuller syndrome

INTRODUCTION

Cleft lip with or without cleft palate and isolated cleft palate (CL/CLP/CP) is a congenital malformation that affects the upper lip, alveolar ridge, tooth eruption, and palate fusion to different degrees. Lip and palate formation is the consequence of several processes that involve cell proliferation, cell differentiation, cell adhesion, and apoptosis. Failure anywhere in these processes can lead to clefts. CL/CLP/CP is one of the most common malformations among newborns (Mooney and Siegel, 2002). Cleft palate (CP) and cleft lip and cleft lip with cleft palate (CL/CLP) are considered etiologically distinct entities, which could be explained by the fact that the lip and palate develop at different embryonic stages (Murray, 2002). The estimated prevalence in the world ranges from 1/300 to 1/2 500 births for CL/CLP and around 1/500 birth for cleft palate only and it varies depending on geographical region and different ethnicities (Stanier and Moore, 2004).

The etiology of non-syndromic CL/CLP/CP is determined by multiple, interacting genetic and environmental factors. Twenty percent of the CL/CLP/CP patients in different populations have a family history of CL/CLP/CP and twin studies showed that proband concordance rate for CL/CLP/CP was 60% in monozygotic (MZ) twins and 10% in dizygotic (DZ) twins, indicating that genetic factors play an important role in the etiology of this birth defect (Murray, 2002). Many genes are considered as susceptibility loci for non-syndromic CL/CLP/CP based on linkage and association studies in different populations. Influence of environmental factors and its interaction with genes involved in embryogenesis also plays a significant role in the CL/CLP/CP development (Stanier and Moore, 2004).

In approximately 30% of the cases CL/CLP/CP is caused by known monogenic syndromes or chromosomal aberrations, and non-syndromic CL/CLP/CP is a complex disease with many contributing genetic factors (Schutte and Murray, 1999). Recent estimates suggest that 2-14 genes could be involved in the formation of CL/CLP/CP (Scliekelman and Slatkin, 2002).

The identification of susceptibility genes for CL/CLP/CP has been the subject of extensive research. To localize candidate genes and loci of non-syndromic clefts, several genome-wide linkage screens, genome-wide association studies and fine mapping have been published. Recent studies have discovered and confirmed regions such as 1p21-p31, 1q32, 2p13, 3q27–28, 4q21-q26, 8q24, 9q21, 10q25.3, 12p11,

14q21–24, 16q24 and 17q22 (Marazita et al., 2004, Riley et al., 2007a, Marazita et al., 2009, Birnbaum et al., 2009, Mangold et al., 2009, 2010). However, despite of the many candidate genes investigated, only the *IRF6* gene has shown a convincing degree of consistency across studies and was considered to be responsible for 12%-18% of non-syndromic CL/CLP/CP cases (Zucchero et al., 2004). These results were replicated in different populations, confirming the role of the *IRF6* gene in CL/CLP/CP formation in different ethnic groups (Marazita et al., 2009). Mutation screening of more than 20 non-syndromic clefts candidate genes showed that only 2%-6% of all screened individuals have mutations in genes including *FOXE1*, *GLI2*, *JAG2*, *LHX8*, *MSX1*, *MSX2*, *SATB2*, *SKI*, *SPRY2*, *TBX10* (Vieira et al., 2005; Jezewski et al., 2003). The recent data suggest that the FGF signaling pathway may contribute to about 3%-5% of non-syndromic CL/CLP/CP cases (Riley et al., 2007b). However other genes studied, such as *TGFA*, *BCL3, PVR*, and *PVRL2* showed conflicting results in genetically diverse populations (Carreno et al., 2002, Pezzetti et al., 2007, Martinelli et al., 1998, Fujita et al., 2004).

Experiments with knockout animal models were conducted to search for new candidate genes for CL/CLP/CP. Few studies with chicks and mice identified specific roles for several major signalling pathways, including Fgf signalling pathways in midfacial morphogenesis and upper lip development (Trumpp et al., 1999). Genetic studies of mice identified two Wnt genes involved in midfacial morphogenesis and CLP development, *WNT3* and *WNT9B* (Juriloff et al., 2001, 2004, 2005, Brugmann et al., 2007).

There is an evidence of marginally increased death rate from cardiovascular disease and cancer in CL/CLP/CP patients. Individuals with non-syndromic CL/CLP/CP have increased death rate from epilepsy, prematurity, pneumonia, aspiration, asphyxia, sepsis and suicide (Christensen et al., 2004).

In nowadays, surgery can repair this defect, but despite this, orofacial clefts have lifelong implications for those affected and their families. That is why there is a necessity for a better understanding of the etiology and the mechanism of cleft formation. Discovering genetic factors involved in the development of CL/CLP/CP will improve the counselling of families at increased risk and will help to predict risk to have an affected offspring. In future identifying environmental factors and its interaction with genetic factors will improve therapy of CL/CLP/CP or even help for prevention.

Aim of the study

The main objective of the study was identification of candidate genes involved in the etiology of non-syndromic cleft lip with or without cleft palate and isolated cleft palate.

Tasks of the study

Fulfillment of the aim required the following tasks:

1. To decide on candidate gene selection in search for significant relationships with non-syndromic cleft lip with or without cleft palate and isolated cleft palate and select genetic markers for further genotyping within the study.
2. To perform case-control association analysis for selected genes to find if genetic variations are assocated with non-syndromic CL/CLP and CP.
3. To perform case-control haplotype analysis for selected genes to find haplotypes with risk or protective effect in the development of non-syndromic CL/CLP and CP, compared to controls.
4. To carry out family-based association analysis for *BCL3*, *PVRL2*, *PVR*, *CLPTM1*, *IRF6* and *BMP4* genes in order to identify transmission distortions.
5. To perform genetic analysis for *BCL3* gene five markers (rs7257231, rs10401176, rs8103315, rs1979377 and rs2927456) for Brazilian non-syndromic cleft with or without cleft palate and isolated cleft palate cases and controls.

Hypothesis of the study

1. Diverse genes and genetic markers are involved in the etiology of non-syndromic cleft lip with or without cleft palate and isolated cleft palate in Latvian population compared to another European origin population.

Scientific novelty of the study

This study is the first study regarding identification of possible candidate genes involved in development of non-syndromic cleft lip with or without cleft palate and isolated cleft palate in Latvian population. Novel finding was *SKI*, *WNT3*, *BMP4*, *IRF6* and *FGFR1* genes role in the development of non-syndromic CL/CLP/CP in Latvian population and obtained results can be used for further studies to identify interaction between genes and environmental factors.

Elaboration of the study

The present study was carried out in the Latvian Biomedical Research and Study Center, Riga, Latvia and Scientific Laboratory of Molecular Genetics, Rīga Stradiņš University, Riga, Latvia in collaboration with University of Pittsburgh, Pittsburgh, USA and University of Tartu, Tartu, Estonia during year 2005-2011.

The Central Medical Ethics Committee of Latvia approved the present study.

The financial support of the study

1. Taiwan-Baltic joint research project No. NSC92-2320-B-075-018. "Identification of genes involved in craniofacial morphogenesis and susceptibility to orofacial clefting in a human genome scan".
2. Latvian Science Council grant No. 06.2021 "Non-syndromic orofacial clefts genetic epidemiological analysis in Latvia".
3. Latvian Science Council grant No. 09.1115 "Risk factor influence on non-syndromic cleft palate, cleft lip with or without palate development in the population of Latvia".
4. ESF project No. 2009/0147/1DP/1.1.2.1.2/09/IPIA/VIAA/009 "Support of the doctoral study program and PhD degree qualification in Rīga Stradiņš University".

Author's contributions

CL/CLP/CP carriers and their parents were recruited from the Riga Cleft Lip and Palate Centre, Institute of Stomatology, Rīga Stradiņš University. Control group was used from Genome Database of Latvian Population and DNA collection of Scientific Laboratory of Molecular Genetics, Rīga Stradiņš University. DNA was extracted at the Scientific Laboratory of Molecular Genetics, Rīga Stradiņš University and Latvian Biomedical Research and Study Centre.

The author performed *COL11A1*, *SKI*, *LHX8*, *IRF6*, *MTHFR*, *TGFA*, *FN1*, *MSX1*, *FGF2*, *NUDT6*, *FGF1*, *MSX2*, *COL11A2*, *EDN1*, *FGFR1*, *FOXE1*, *TBX10*, *MMP3*, *MMP13*, *PVRL1*, *COL2A1*, *SPRY2*, *BMP4*, *TGFB3*, *JAG2*, *MMP25*, *MMP2*, *CDH1*, *RARA*, *WNT3*, *WNT9B*, *TIMP2*, *SMAD2*, *SMAD4*, *BCL3*, *PVRL2*, *PVR*, *CLPTM1*, *APOC2*, *BMP2*, *MMP9*, *TIMP3*, *SYN3*, *TBX22* and *TIMP1* genes genotyping.

All data statistical analysis was performed by the author of this thesis.

Outline of the thesis

The thesis is composed on 166 pages in English, following classical scheme, structured in ten chapters: Introduction, Review of literature, Subjects and Methods, Results, Discussion, Conclusions, Publications, Acknowledgements, References and Appendixes. Text of thesis is supplemented by 3 figures, 50 tables and 12 appendixes. Reference list consist of 139 cited references.

1. REVIEW OF LITERATURE

1.1. Classification of lip, alveolar ridge and palate clefts

The most common classification is dividing deformity into cleft lip, cleft lip with cleft palate and isolated cleft palate. Cleft lip can be divided into unilateral and bilateral cleft lip. In the unilateral cleft lip case nasolabial and bilabial muscle rings are disrupted on one side, and it is resulting in asymmetrical deformity involving external nasal cartilage, nasal septum and maxilla. Unilateral clefts can be divided into left and right side clefts. Bilateral cleft lip means that two muscular rings are disrupted on both sides and it is resulting in symmetrical deformity. Cleft palate can be classified as incomplete and complete. Incomplete cleft means when the cleft of the hard palate remains attached to the nasal septum and vomer, but in the complete cleft case the nasal septum and the vomer are completely seperated from the palatine process (http://www.who.int) (see Figure 1).

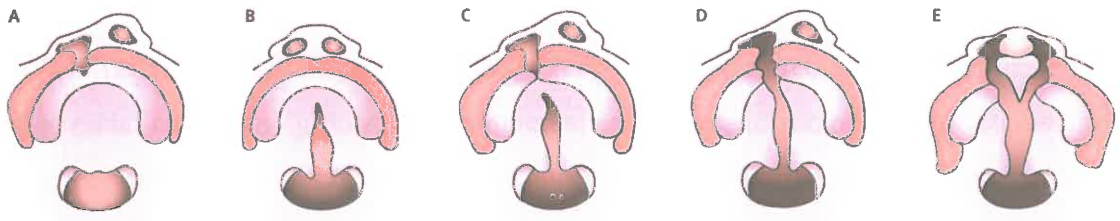

Figure 1. Classification of cleft lip with or without cleft palate and isolated cleft palate

A - Cleft lip and alveolus. B - Cleft palate. C - Incomplete unilateral cleft lip and palate. D - Complete unilateral cleft lip and palate. E - Complete bilateral cleft lip and palate. Adapted from Mossey et al., 2009; Shaw W.C. Orthodontics and occlusal management. Oxford: Butterworth-Heinemann, 1993.

Recenty used is LAHSHAL (L - lip, A - alveolus, H - hard palate, S - soft palate, H - hard palate, A - alveolus, L - lip) classification (Kriens, 1987), which describes site, size, extent and type of cleft. List of LAHSHAL codes are shown in Appendix 1.

All orofacial clefts can be divided into syndromal and non-syndromal forms. Chromosomal aberrations, more than 600 different recognizable monogenic syndromes (www.ncbi.nlm.nih.gov/omim/), teratogen-induced disorders and also unrecognized syndromes form syndromic type of CL/CLP/CP (Schutte and Murray, 1999; Christensen, 2004, Dixon et al., 2011).

The term "non-syndromic" is restricted to CL/CLP/CP where the afftected individuals have no other physical or developmental anomalies, no recognized maternal environmental exposures, no chromosomal aberrations and monogenic condition (Murray, 1995).

Currently most studies suggest that in human, approximately 70% of all cases of CL/CLP/CP and 50% of isolated cleft palate cases are considered to be non-syndromic (Jones, 1988; FitzPatrik and Farrall, 1993; Marazita 2002).

Cleft lip and cleft lip with cleft palate are categorized together because these two phenotypes are thought to have the same genetic etiology, whereas isolated cleft palate have different genetic background (Harville, 2005). Despite this recent studies have found evidence that cleft lip and cleft lip with cleft palate might be separate entities with different etiology and pathogenesis (Jugessur et al., 2011).

1.2. Epidemiology of cleft lip with or without cleft palate and isolated cleft palate

CL/CLP/CP affects approximately 1/700 of live borns, with wide variability across racial and ethnic groups. Environmental exposure and socioeconomic status also influence prevalence of CL/CLP/CP. In general, Asian and native American populations have the highest birth prevalence rates for CL/CLP/CP, often as high as 1/500, European-ancestry populations have intermediate prevalence rates at about 1/1000, and African-ancestry populations have the lowest prevalence rates at about 1/2500 of live births. These observations are suggestive of relative individual contribution susceptibility genes, which may vary between different populations (Mossey, 2009). CL/CLP/CP prevalence in Latvia and Lithuania is ~ 1/700 and, according to personal communication, similar in Estonia (Akota et al., 2001; Morkuniene et al., 2007). Figure 2 shows birth prevalence of non-syndromic cleft lip with or without cleft palate and isolated cleft palate in Europe.

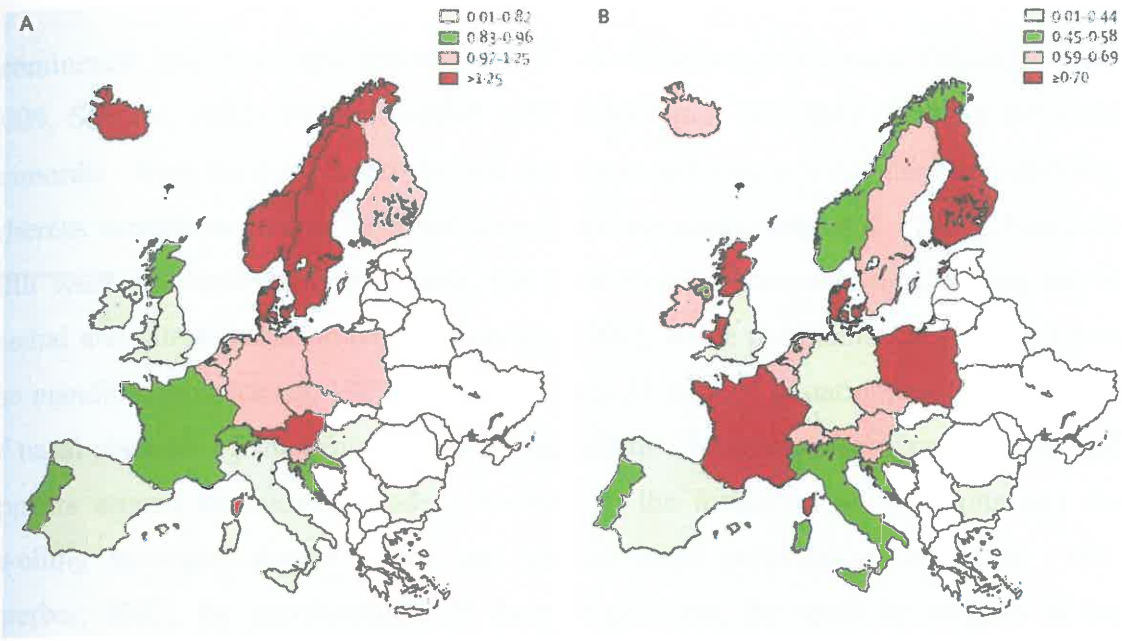

Figure 2. Birth prevalence per 1000 live births of non-syndromic cleft lip with or without cleft palate and isolated cleft palate in Europe
A - Cleft lip and cleft lip with cleft palate (CL/CLP). B - Cleft palate only (CP). Adapted from Mossey et al., 2009 and http://www.eurocran.org.

The prevalence of CL/CLP/CP also differs by sex and laterality. There is a 2:1 male to female ratio of CL/CLP patients and approximately a 1:2 male to female ratio of CP patients. There is a 2:1 ratio of left side clefts prevalence among CL/CLP cases (Dixon et al., 2011).

1.3. Craniofacial development

Lip and palate formations are the consequence of several processes that involve cell proliferation, cell differentiation, cell adhesion, and apoptosis. Impairment in any of these processes can lead to CL/CLP/CP. Genetic variations interacting with environmental factors, convergence and fusion of the facial and palatal processes, apoptosis, and adequate nutrient supply can alter normal craniofacial development, but can also contribute to abnormal lip and palate development (Mossey et al., 2009).

1.3.1. Normal upper lip development

Normal development of the human face begins in the fourth week of embryogenesis, when neural crest cells migrate through mesenchymal tissue into the developing craniofacial region. These cells participate in formation of the frontonasal

prominence, two maxillary processes and two mandibular processes (Mossey et al., 2009, Sperber, 2002). In combination with mesodermal cells they establish the facial primordia - from the neural crest-derived facial mesenchyme will rise the facial skeleton, whereas mesoderm-derived cells will form facial muscles (Jiang et al., 2006). Fourth to fifth week of human embryogenesis, the frontonasal prominence divides into paired medial and lateral nasal processes (Sperber, 2002), while the manibular process forms the mandible (Hinrichsen, 1985). Approximately 32 days of gestation, occurs formation of nasal placodes by thickening of surface ectoderm. The frontonasal process grows and appears around the nasal placodes, resulting in the formation of nasal pits and the swelling horseshoe-shaped lateral and medial nasal processes (Hinrichsen, 1985; Sperber, 2002). By approximately 35 days of gestation, the upper lip consists of the maxillary processes laterally and the medial nasal processes medially with the lateral nasal processes wedged in between the medial nasal and maxillary processes. Fusion between the medial and lateral nasal processes has initiated while maxillary processes lie below the lateral nasal processes. By approximately 38 days of gestation in human, the maxillary and medial nasal processes grow rapidly and push the lateral nasal processes further and bring the distal ends of maxillary and medial nasal processes into direct contact. By the end of the 6th week of development, merging of the medial nasal processes with each other and with the maxillary processes on each side leads in formation of the upper lip (Sperber, 2002).

1.3.2. Normal palate development

Palate development begins during the 5th week of embryogenesis after fusion of the upper lip and by the end of the 6th week primary palate is formated. Secondary palate starts to develop during the 6th week with outgrowth from the maxillary processes of two palatal shelves, which initially grow vertically down the sides of the developing tongue. During the 7th week of embryogenesis, the palatal shelves rise to a horizontal position above the tongue and come into contact and fuse. The palatal mesenchyme differentiates into bony and muscular elements which are correlated with the position of the hard and soft palate. Afterwards the secondary palate fuses with primary palate and the nasal septum. Development and fusion of secondary palate is finished by the 10th week of embryogenesis (Mossey et al., 2009).

1.3.3. Development of cleft lip with or without cleft palate and isolated cleft palate

Cleft lip and palate is the result of improper fusion of the processes that form the face, caused by abnormal morphogenesis of the upper lip and primary palate either by misguided epithelial movement, disrupted epithelial-mesenchymal transformation (EMT), or disrupted apoptosis (Jiang et al., 2006). Failure of these mechanisms result in insufficient growth, decreased nutrients and/or a diminished degradation of the epithelial seam covering the growth processes, each predisposing cleft lip and palate. A cleft of the lip and/or palate results from a failure of fusion of the frontonasal prominence with the medial nasal process, the primary palate with the secondary palate or the lateral palatine shelves with each other (Nanci, 2003).

Since fusion between the secondary palatal shelves, which arise bilaterally from the maxillary processes (Ferguson, 1988), and fusion between the primary and secondary palates occur much later in embryogenesis than the fusions between maxillary, lateral and medial nasal processes during lip formation, failure of proper lip fusion often affects palatal contact secondarily. Therefore, cleft lip is often accompanied by cleft palate (Jiang et al., 2006).

D.G. Trasler (1968) emphasized the importance of fusion between medial and lateral nasal processes and postulated that lateral cleft lip results when this fusion process does not occur.

1.4. Etiology of non-syndromic cleft lip with or without cleft palate and isolated cleft palate

P. Fogh-Andersen (1942) first collected data of families with non-syndromic clefts and evaluated the observed pattern of inheritance. He concluded that the families with CL/CLP were consistent with segregation of alleles at a single genetic locus with variable penetrance, but families with CP were consistent with autosomal dominant inheritance with reduced penetrance. At 1965 D.S. Falconer developed specific statistical model of inheritance called "multifactorial treshold model" (MFT) in order to explain the familial patterns of CL/CLP/CP and in 1969 C.O. Carter suggested that the familial aggregation patterns could be explained by the MFT model and this model at that time was considered to be the most appropriate model of inheritance for CL/CLP/CP. F.C. Fraser (1976) postulated that the common congenital malformations

have familial distributions that cannot be accounted for by simple Mendelian models, but can be explained in terms of a continuous variable, "liability," with a threshold value beyond which individuals will be affected. Both genetic and environmental factors determine liability, making the system multifactorial. The term "multifactorial" should be used for "determined by a combination of genetic and environmental factors," without reference to the nature of the genetic factor(s). "Polygenic" should be reserved for "a large number of genes, each with a small effect, acting additively." When several genes, with more major effects are involved, "multilocal" can be used. When it is not clear which of these is applicable the term "plurilocal" is suggested, in the sense of "genetic variation more complex than a simple Mendelian difference" (Fraser, 1976). Few years later M. Melnick et al. (1980) concluded that observed data does not provide strong evidence in favor of MFT inheritance. He tested MFT as a inheritance model and it revealed that the incidence of CL/CLP/CP in siblings was 40 times greater than that in the general population; the risk to siblings of CL/CLP/CP females was not significantly different from the risk to siblings of CL/CLP/CP males; recurrence risk for siblings of CL/CLP/CP probands was dependent upon the proband's cleft type; only 0.4% of the variation in risk to the siblings born after the proband could be accounted for by the number of previously affected siblings; the consanguinity rate was six times less than the general population rate; heritability estimates from siblings and parents by sex suggest, either the presence of significant dominance effects, or a common sibling environment component in the etiology of the malformation (Melnick et al., 1980). In addition, the results of several segregation analyses have been interpreted as providing strong evidence in favor of major-gene effect in the etiology of non-syndromic CL/CLP/CP. In most published segregation analysis of CL/CLP/CP rejected MFT model and in favor of a mixed model (single major locus plus multifactorial components) (Marazita et al., 1984, 1986; Chung et al., 1986) or a major locus alone (Hecht et al., 1991). Analysis of recurrence-risk patterns showed another model of inheritance named oligenic or multilocus model of inheritance with approximately four to seven interacting loci (Mitchell and Risch, 1992; Christensen and Mitchell, 1996). Several twin studies were performed to establish mode of inheritance for CL/CLP/CP. Obtained results showed that the proband concordance rate for CL/CLP/CP was 60% in monozygotic twins (MZ) and 10% in dizygotic twins (DZ). This finding indicates that genetic factors play a role in the cause of CL/CLP/CP, but environmental factors are probably involved too (Shields et al., 1979; Christensen and Fogh-Andersen, 1993).

Murray (2002) reported that lack of 100% concordance in monozygotic twins suggests that genetic factors alone are not responsible for cleft phenotype and greatly increased MZ concordance strongly supports a major genetic component. Additional genetic linkage and association studies are used to identify these genetic factors. It is proved now that environmental factors have influence in the development of non-syndromic cleft lip and palate (Murray, 2002, Mossey, 2009; Wehby and Murray, 2010). For now a multifactorial model of inheritance is favored in which genetic risk factors of small individual impact may interact with environmental factors (Rahimov et al., 2008). These combined factors complicate genetic analysis of non-syndromic forms of CL/CLP/CP (Dixon et al., 2011).

1.5. Genetic factors

There are many different genetic approaches such as genome-wide and few genes linkage scans, genome-wide and few genes association studies, fine mapping, informative mouse models and also gene expression studies in mouse and human embryonic tissues to identify new genes involved in the etiology of CL/CLP/CP and clarify the role for previously reported genes. Recent studies discovered and confirmed chromosomal regions such as 1p21-p31, 1q32, 2p13, 3q27–28, 4q21-q26, 8q24, 9q21, 10q25.3, 12p11, 14q21–24, 16q24 and 17q22 (Marazita et al., 2004, Riley et al., 2007a, Marazita et al., 2009, Birnbaum et al., 2009, Mangold et al., 2009, 2010). However, despite of the many candidate genes investigated, only the *IRF6* gene has shown a convincing degree of consistency across studies and was considered to be responsible for 12%-18% of non-syndromic cleft lip with or without cleft palate cases (Zucchero et al., 2004). These results were successfully approved in unrelated populations from Italy, Norway, Belgium, the USA, Thailand, South America, and China, confirming the role of the *IRF6* gene in CL/CLP/CP formation in different ethnic groups (Marazita et al., 2009). Mutation screening of more than 20 non-syndromic cleft lip with or without cleft palate candidate genes showed that only 2%-6% of all screened individuals have mutations in the rest of genes including *FOXE1*, *GLI2*, *JAG2*, *LHX8*, *MSX1*, *MSX2*, *SATB2*, *SKI*, *SPRY2*, *TBX10* (Vieira et al., 2005; Jezewski et al., 2003). The recent data suggest that the *FGF* signaling pathway may contribute to about 3%-5% of non-syndromic CL/CLP/CP cases (Riley et al., 2007b). These results indicate that, besides *IRF6*, there are many more genes involved in the etiology of non-syndromic

CL/CLP/CP. However other genes studied, such as *TGFA*, *BCL3, PVR*, and *PVRL2* showed conflicting results and failed replication studies in genetically diverse populations (Carreno et al., 2002, Pezzetti et al., 2007, Martinelli et al., 1998, Fujita et al., 2004).

1.5.1. FGFs and their receptors (FGFRs)

The human fibroblast growth factors (FGFs) are encoded by 18 distinct genes (*FGF1-FGF10* and *FGF16-FGF23*), that play pleiotropic roles in human development and metabolism. The biological activities of the FGFs are mediated by FGF receptor tyrosine kinases (FGFRs) encoded by four distinct genes (*FGFR1–FGFR4)* in mammals (Turner and Grose, 2010). FGF signalling controls cell proliferation, migration, differentiation, survival, and thus plays essential roles in various processes of embryonic development (Thisse and Thisse, 2005). Abnormal FGF-FGFR signaling due to gain- or loss-of- function mutations or misexpression has been implicated in a multitude of human diseases, including craniosynostosis, chondrodysplasia and Kallmann syndromes, cancers, and isolated hypogonadotropic hypogonadism (IHH) (Dode et al., 2003; Marie et al., 2005; Trarbach et al., 2007; Turner and Grose, 2010).

1.5.2. *SKI* gene

SKI (v-ski sarcoma viral oncogene homolog (avian)) gene, located in 1q22-q24, was discovered as an oncogene present in the avian Sloan-Kettering viruses. It was most likely formed during the passage of a transformation-defective avian leucosis virus, which was derived from a cellular gene, c-SKI, which is proto-oncogene (Stavnezer, 1989; Sutrave, 1989). The gene encodes a nuclear protein that binds to DNA and modulates transcription in association with other cellular factors known to be involved in craniofacial development, including *SMAD2*, *SMAD3*, and *SMAD4* (SMAD family memeber 2, 3 and 4) (Nagase, 1990; Engert, 1995; Berk, 1997). M. Berk et al. (1997) originally reported in their study that the c-SKI proto-oncogene has been implicated in the control of cell growth and skeletal muscle differentiation. To determine its normal functions in vivo, authors disrupted the mouse c-SKI gene. The results show a novel role for *SKI* gene in the morphogenesis of craniofacial structures and the central nervous system, and confirm its proposed function as a player in skeletal muscle development.

Homozygous mutant mice show perinatal lethality resulting from exencephaly, a defect caused by failed closure of the cranial neural tube during neurulation (Berk et al., 1997).

Ski protein is an important negative regulator of the Smad proteins. Ski can bind to the BMP-Smad protein complexes in response to BMP and repress their ability to activate BMP target genes through disruption of a functional Smad complex and through recruitment of transcriptional co-repressors. The antagonism of BMP signaling by Ski results in neural specification in Xenopus embryos and inhibition of osteoblast differentiation in mouse bone-marrow stromal progenitor cells. This ability to modulate BMP signaling by Ski may play an important role in the regulation of craniofacial, neuronal, and skeletal muscle development (Luo et al., 2003).

It has been also suggested that *SKI* gene is involved in the development of palate by the cAMP and TGFB signalling pathways (Warner et al., 2003).

1.5.3. *WNT* family genes

The *WNT* gene family (wingless-type MMTV integration site family) consists of structurally related genes that encode cysteine-rich secreted glycoproteins that act as extracellular signaling factors. Because of their role in the regulation of cell fate and patterning during embryogenesis, including craniofacial development, members of this family are biologically important candidates for non-syndromic and syndromic clefts in humans (Gavin et al., 1990; Chiquet et al., 2008).

A role for Wnt signaling in facial morphogenesis was not known until the identification of the *WNT3* (wingless-type MMTV integration site family, member 3) nonsense mutation in humans. *WNT3* gene, located at 17q21, is required at the earliest stages of human limb formation and for craniofacial and urogenital development (Niemann et al., 2004). Genetic analysis results of seven *WNT* family genes suggest that alteration in Wnt gene function may perturb formation or fusion of the facial processes and predispose carriers to CL/CLP/CP (Chiquet et al., 2008).

1.5.4. Collagen family genes

Collagens II and XI are present throughout Meckel's cartilage, which provides mechanical support for the developing mandible (Chung et al., 1995) and this is one of reasons why these genes can be considered as candidate genes for non-syndromic clefts.

Mutations in collagen genes are involved in different syndromes. Recessive mutations in *COL11A2* (collagen, type XI, alpha 2) are responsible for otospondylomegaepiphyseal dysplasia (OSMED) and non-syndromic hearing loss while dominant mutations are associated with Stickler type III, isolated cleft palate, Robin sequence, non-ophthalmic Stickler syndrome, early onset osteoarthritis and autosomal dominant hearing loss (Kayserili et al., 2011). Stickler syndrome caused by mutations in *COL2A1*, *COL11A1*, or *COL11A2* is inherited in an autosomal dominant manner, but Stickler syndrome caused by mutations in *COL9A1* or *COL9A2* is inherited in an autosomal recessive manner (Robin et al., 2000). Other phenotypes associated with mutations in *COL11A2* is Weissenbach-Zweymuller syndrome (WZS) and non-syndromic sensorineural hearing loss (van Steensel et al., 1997). It has been considered that collagen can be involved only in syndromic clefts, however there is a study which shows that *COL2A1*, *COL11A1* and *COL11A2* can be involved in the etiology for non-syndromic CL/CLP/CP (Melkoniemi et al., 2003).

1.5.5. 19q13 locus

Locus 19q13 has been suggested as a susceptibility region for cleft development. The *BCL3* gene is localized in chromosome 19q13 in close proximity to genes previously associated with cleft phenotypes such as *PVR*, *PVRL2* and *CLPTM1* (Warrington et al., 2006), and encodes a transcription factor involved in cell cycle regulation. Hence, *BCL3* may be involved in lip and palate morphogenesis for its role in mediating cell differentiation, and other cell processes during embryonic development (McKeithan et al., 1987).

Stein et al. (1995) presented the first evidence that *BCL3* plays a role in the etiology of non-syndromic CL/CLP/CP through linkage analysis, assuming an autosomal dominant type of inheritance with incomplete penetrance (Stein et al., 1995). Other studies have also found linkage between the *BCL3* locus and non-syndromic CL/CLP/CP (Amos et al., 1996a; 1996b; Wyszynski et al., 1997). Maestri et al. (1997) reported a significant association between *BCL3* and CL probands (Maestri et al., 1997). In addition, Martinelli et al. (1998) found marginal association between a highly polymorphic intragenic marker (D19S574) close to *BCL3* in 98 infants with CL/CLP and their parents (Martinelli et al., 1998). Beaty et al. (2001) also observed an excess transmission of this same marker allele, although no formal statistical difference was

found between allele frequencies in CP cases and controls. Studies in Chilean populations also showed a small difference in *BCL3* allele distribution between non-syndromic CL/CLP cases and controls (Blanco et al., 2004; Carreno et al., 2002). However, no evidence of linkage was found between *BCL3* in non-syndromic CL/CLP Japanese families (Fujita et al., 2004), likewise no association in Lithuanian population (Morkuniene et al., 2007). Finally, it has been suggested that *BCL3* plays a role in the etiology of non-syndromic CL/CLP as a low penetrance gene or as a modifier (Gaspar et al., 2002).

1.5.6. *BMP4* gene

BMP4 (bone morphogenic protein 4) gene, located at 14q22-q23 in humans, is a member of the transforming growth factor beta superfamily. Expression studies of bone morphogenic proteins (BMPs) and its antagonist Noggin in the embryonic chicken face suggested that BMP signals were important for closure of the upper lip or primary palate (Ashique et al., 2002). The same authors performed gain- and loss-of-function experiments to determine the role of BMPs in lip formation and they found that BMPs regulate outgrowth and epithelial survival during avian lip fusion. Liu et al. (2005) presented that conditional inactivation of Bmp4 in a transgenic mice line results in an isolated cleft lip. Because of its role in the regulation of skeletal development including cartilage and bone formation during craniofacial and limb development *BMP4* has been suggested as candidate gene for non-syndromic CL/CLP/CP (Wan and Cao, 2006).

1.5.7. *IRF6* gene

IRF6 gene, located in 1q31-41, is suggested as only one gene, which involvement in the development of CL/CLP/CP is confirmed in many unrelated populations (Marazita et al., 2009). High levels of *IRF6* mRNA have been discovered along the medial edge of the fusing palate and also tooth buds (Kondo et al., 2002). Author demonstrated that haploinsufficiency of *IRF6* can lead to cleft lip with or without cleft palate and lack of teeth. T.M. Zucchero et al. (2004) found very strong association between V274I polymorphism of *IRF6* and non-syndromic CL/CLP/CP in different populations and this study was used as a background for further studies to clear role of

IRF6 gene in etiology of non-syndromic clefts compared between different populations (Zucchero et al., 2004).

1.6. Environmental factors

Identification of environmental components of clefting and studies of gene by environment interaction require large and in the best case prospective cohort studies and access to genetic material to be optimally effective (Dixon et al., 2011).

Maternal smoking has been associated with increased risk of CL/CLP/CP and meta-analysis strongly supports an overall odds ratio (OR) for having CL/CLP/CP of ~1.3 among offspring of mothers who smoke (Little et al., 2002; Shi et al., 2007; Shi et al., 2008). Increased risks from exposure to maternal smoking during the peri-conceptual period raises the possibility that genes in certain metabolic pathways may play a role in the development of CL/CLP/CP. Specifically, markers in the *GSTT1* (glutathione S-transferase theta 1) or *NOS3* (nitric oxide synthase 3) genes appear to influence risk of CL/CLP/CP in the presence of maternal smoking (Shi et al., 2007; von Rooij et al., 2001; Lammer et al., 2004; Zhu et al., 2009). The *GSTT1* markers are gene deletion variants, which suggest deficiencies in detoxification pathways may underlie some of this susceptibility. Smoking has also been recently associated with a joint risk with variants in the *IRF6* gene and the same study reported interactions between multivitamins and *IRF6* variants (Wu et al., 2010). These findings provide evidence that gene-environment interactions are important in CL/CLP/CP.

Nutritional factors, such as folate deficiency, have also been suggested to influence risk of CL/CLP/CP, based on both observational studies and interventional trials using folate supplementation to prevent recurrences of CL/CLP/CP/ in families (Wehby and Murray, 2010). However, the studies of vitamin supplementation with folate remain controversial (Wilcox et al., 2007; Wehby and Cassell, 2010) and recent studies of levels of folate receptor antibodies did not find an association with CL/CLP/CP (Bille et al., 2010). Furthermore, food fortification programs using folic acid have shown detectable decreases in the rates of clefting in some (Yazdy et al., 2007; Johnson and Little, 2008), but not all studies (Ray et al., 2003; Lopez-Camelo et al., 2010). There are studies which support roles for zinc deficiency in risk of CL/CLP/CP (Munger et al., 2009), for cholesterol deficiency in CL/CLP/CP (Porter, 2006) as well for as multivitamins in general in CL/CLP/CP prevention (Johnson and

Little, 2008), but additional studies have to be made to confirm it (Dixon et al., 2011).

In addition, some specific teratogens, for example valproic acid, have yielded evidence of association with cleft palate (Jentink et al., 2010).

Exposure to maternal alcohol consumption has also been suggested as a risk factor, but the evidence has been more inconsistent (Mossey, 2009). Studies also suggest that high doses of alcohol in short periods of time increase risk (Deroo et al., 2008), and this is supported by associations with variation in the *ADH1C* alcohol dehydrogenase gene (Boyles et al., 2009).

Besides nutrients and toxins there are other environmental factors such as hyperthermia, stress, maternal obesity, occupational exposures, ionizing radiation and infection, but the effects of these factors are not yet clarified (Shahrukh et al., 2010; Mossey et al., 2009).

2. SUBJECTS AND METHODS

2.1. Subjects

In the study were included patients with non-syndromic cleft lip (CL), patients with non-syndromic cleft lip with cleft palate (CLP), patients with non-syndromic cleft palate (CP), patients with no age or sex limit and from Caucasian origin. Patients with syndromic clefts or any recognized inherited pathology and adopted patients were excluded from the study. Control group consisted from unrelated, randomly selected unaffected individuals with no family history of clefts, with no age or sex limit and from Caucasian origin. Individuals with syndromic clefts or any recognized inherited pathology and adopted individuals were excluded from the study.

The data collection was performed in accordance with the regulations issued by the Central Medical Ethics Committee of Latvia. Prior to any research procedure, all participated individuals signed an informed consent form. In the case of patients who were under 18 years of age, consent was obtained from their parents.

Patients and their parents were recruited in the Riga Cleft Lip and Palate Centre, Institute of Stomatology, Rīga Stradiņš University.

The control group consisted of 190 individuals collected at the Latvian Biomedical Research and Study Center within the framework of the national project Genome Database of Latvian Population and 293 individuals from internal data collection of Scientific Laboratory of Molecular Genetics, Rīga Stradiņš University.

For case-control study 661 individual from Latvia were analyzed: 178 non-syndromic cleft lip with or without cleft palate and isolated cleft palate (CL/CLP/CP) patients and 483 unaffected individuals as controls. Out of all 178 non-syndromic CL/CLP/CP cases, 135 had CL/CLP (36 patients with CL, 99 patients with CLP) and 43 patients had CP.

The transmission disequilibrium test (TDT) was carried out in Latvian 122 trios (patient with both parents, total 366 persons), from which 89 patients and their parents (total 267 persons) were divided into CL/CLP group and 33 patients with both parents (total 99 persons) - into CP group.

For additional study 606 DNA samples from Brazilian population (338 non-syndromic CL/CLP/CP cases and 268 individuals as controls) were studied in present study. In the study were included patients with non-syndromic cleft lip (CL), patients

with non-syndromic cleft lip with cleft palate (CLP), patients with non-syndromic cleft palate (CP), patients with no age or sex limit and from Caucasian origin. Patients with syndromic clefts or any recognized inherited pathology were excluded from the study. Control group consisted from unrelated, randomly selected unaffected individuals with no family history of clefts, with no age or sex limit and from European origin (Portugese descent). Individuals with syndromic clefts or any recognized inherited pathology were excluded from the study. Out of all 338 cleft cases, 294 cases had non-syndromic CL/CLP and 44 cases had non-syndromic CP. Brazilian population samples were obtained at the Dental Clinics of the Hospital of Rehabilitation and Craniofacial Anomalies and Bauru Dental School, both of the University of São Paulo. Bauru, SP, Brazil. The study had local approval and was conducted with the consent of the participants and their parents or legal guardians.

2.2. Methods

2.2.1. DNA extraction

The genomic DNA of non-syndromic CL/CLP/CP patients and population samples was obtained from venous blood or saliva and extracted according to the established protocol of the phenol-chloroform method with slight modifications.

2.2.1.1. DNA extraction and purification from venous blood I

1. Add 2-5 ml of Lysis buffer and 60-120 µl of TWEEN 20 to 2-5 ml of venous blood and mix slowly using rotator Multi RS-60 (*Biosan*, Latvia) 10 minutes at room temperature following by spin for 8 minutes at 3000 rpm/min.
2. Discard the supernatant.
3. Add 700 µl of 10 mM Tris-HCl, 10 mM EDTA, 0.456 M NaCl solution, 100 µl 1.6% SDS and 6 µl proteinase K (*Fermentas*, Lithuania).
4. Incubate at 37^0C for overnight or 55^0C for 2 hours.
5. Transfer all solution into new tube and add 400 µl buffered phenol, pH 8.0 and mix the tube slowly using rotator Multi RS-60 (*Biosan*, Latvia) for 10 minutes at room temperature following by spin for 3 minutes at 15 000 rpm/min.
6. Transfer the supernatant to a new tube and add 200 µl buffered phenol, pH 8.0 and 200 µl chlorophorm-isoamylalcohol solution (in ratio 24:1).

7. Mix the solution slowly using rotator Multi RS-60 (*Biosan*, Latvia) for 10 minutes at room temperature following by spin for 3 minutes at 15 000 rpm/min.
8. Transfer the supernatant into new tube, add cold 96% ethanol in ratio 1:1 and mix gently.
9. Spin the solution for 3 minutes at 15 000 rpm/min.
10. Discard the supernatant and dry the tube for 20 minutes with open cover.
11. Add 200 µl TE buffer or dH_2O.
12. DNA sample storage performed at -20^0C.
13. DNA concentration measurement performed with NanoDrop 2000 Spectrophotometer (*Thermo Fisher Scientific*, USA).

2.2.1.2. DNA extraction and purification from white blood cells II

1. Spin tube with venous blood at $+4^0$C for 15 minutes at 4000 rpm/min.
2. Prepare 50 ml tube and add ~ 10 ml RBC Lysis A1 solution.
3. Transfer leukocytes from blood tube to prepared 50 ml tube, add once more ~ 10 ml RBC Lysis A1 solution and mix gently by inverting.
4. Incubate at $+4^0$C for 15 minutes.
5. Spin tube at $+4^0$C for 15 minutes at 4000 rpm/min.
6. Discard the supernatant and mix gently the suspension.
7. Transfer suspension to a new 15 ml tube and add 5 ml Cell Suspension Solution.
8. Mix the solution slowly using rotator Multi RS-60 (*Biosan*, Latvia) for 3-5 minutes at room temperature.
9. Add 400 µl 10% SDS solution and invert the tube ~ 3 times.
10. Add 5 µl proteinase K and invert the tube ~ 3 times.
11. Incubate at $+50^0$C for overnight.
12. Add 5 ml buffered phenol and mix the tube slowly using rotator Multi RS-60 (*Biosan*, Latvia) for 15 minutes at room temperature following by spin at $+20^0$C for 10 minutes at 4000 rpm/min.
13. Transfer the upper level to a new tube and add 5 ml chloroform.
14. Mix the tube slowly using rotator Multi RS-60 (*Biosan*, Latvia) for 5 minutes at room temperature following by spin at $+20^0$C for 10 minutes at 4000 rpm/min.
15. Transfer the upper level to a new tube and add very careful 5 ml isoamylalcohol.

16. Invert the tube few times following by spin at +20^{0}C for 10 minutes at 4000 rpm/min.
17. Discard the supernatant and add 5 ml 70% ethanol.
18. Vortex the tube for 10 seconds and incubate it for 2 minutes at room temperature.
19. Spin at +20^{0}C for 10 minutes at 4000 rpm/min.
20. Discard the ethanol and dry the tube for 10 minutes at room temperature.
21. Add 1 ml TE buffer and mix gently using rotator Multi RS-60 (*Biosan*, Latvia) for overnight at room temperature.
22. DNA sample storage performed at -20^{0}C.
23. DNA concentration measurement performed with NanoDrop ND-1000 Spectrophotometer (*Thermo Scientific*, USA).

2.2.1.3. DNA extraction from dried blood samples

1. Cut 3 mm (in diameter) of dried blood sample and put it into 1.5 ml tube.
2. Add 1 ml ddH_2O and put into rotator for 30 min.
3. Remove ddH_2O by pipette.
4. Add 100 µl of methanol.
5. Incubate 15 min at room temperature (tube is closed).
6. Remove methanol by pipette.
7. Add 100 µl of 5 mM NaOH and few drops of mineral oil.
8. Incubate 10 min at 100°C.
9. Put tube on ice and incubate 10-15 min.
10. Store DNA at -20°C.
11. DNA concentration measurement performed with NanoDrop ND-1000 Spectrophotometer (*Thermo Scientific*, USA).

2.2.1.4. DNA extraction and purification from saliva

1. Add 2 ml SLS-1 solution to 1-2 ml saliva and mix it.
2. Add 5 µl proteinase K (*Fermentas*, Lithuania) and invert the tube ~ 3 times.
3. Incubate overnight at +50^{0}C.

4. Add 3-4 ml buffered phenol, mix the tube slowly using rotator Multi RS-60 (*Biosan*, Latvia) for 15 minutes at room temperature following by spin at +20^{0}C for 10 minutes at 4000 rpm/min.
5. Transfer the upper level (~ 2-2.5 ml) to a new tube and add 2-2.5 ml chloroform.
6. Mix the tube slowly using rotator Multi RS-60 (*Biosan*, Latvia) for 5 minutes at room temperature following by spin at +20^{0}C for 10 minutes at 4000 rpm/min.
7. Transfer the upper level (~ 2-2.5 ml) to a new tube and add very carefully 2-2.5 ml isoamylalcohol.
8. Invert the tube until the DNA pellet forms.
9. Spin at +20^{0}C for 10 minutes at 4000 rpm/min.
10. Discard all supernatant, add 5 ml 70% ethanol and vortex for 10 seconds following by mix using rotator Multi RS-60 (*Biosan*, Latvia) for 5 minutes at room temperature.
11. Spin at +20^{0}C for 10 minutes at 4000 rpm/min and discard the ethanol immediately.
12. Dry the tube for 10 minutes at room temperature.
13. Add 250 µl TE buffer, vortex briefly and mix gently using rotator Multi RS-60 (*Biosan*, Latvia) for overnight at room temperature.
14. DNA sample storage performed at -20^{0}C.
15. DNA concentration measurement performed with NanoDrop ND-1000 Spectrophotometer (*Thermo Fisher Scientific*, USA).

2.2.2. Genotyping

2.2.2.1. APEX-2 (Arrayed primer extension reaction) method

Arrayed primer extension reaction is a straightforward and robust enzymatic genotyping method in which hundreds to thousands of variations in the genome are simultaneously analyzed in a single multiplexed reaction. APEX occurs by a two-step reaction mechanism: (1) targeting of DNA hybridization to the complementary oligo-primers and (2) single base extension of these primers with appropriate dye- labeled dideoxynucleotides that match the nucleotide on polymorphic site by DNA polymerase. Thus, these dye-labeled dideoxynucleotides are used to terminate the extension reaction

directly at their incorporation site, complementarily representing the DNA base in question (Pullat and Metspalu, 2008).

2.2.2.1.1. Gene and SNP selection

To capture all of the SNPs with minor allele frequencies MAF > 0.05 and r^2 = 0.8 in the regions of interest, 651 tagSNPs were selected based on the HapMap Phase II data, using HapMap CEU as a reference population. Multiple SNPs were selected for each gene, including 10 kb of both upstream and downstream genomic sequences.

A list of selected SNPs per gene are shown in Table 2.1.

Table 2.1.

Candidate genes and loci included in the study

Gene/Locus	Chromosomal localization	Number of genotyped SNPs^
MTHFR	1p36.3	11
LHX8	1p31.1	9
COL11A1	1p21	48
SKI	1q22-q24	20
IRF6	1q32.3-q41	11
TGFA	2p13	41
FN1	2q34	30
MSX1	4p16.3-p16.1	15
FGF2	4q26-q27*	20
FGF1	5q31	35
MSX2	5q34-q35	6
EDN1	6p24.1	15
COL11A2	6p21.3	22
FGFR1	8p11.2-p11.1	12
FOXE1	9q22	4
TBX10	11q13.2	10
MMP3	11q22.3	8
MMP13	11q22.3	20
PVRL1	11q23.3	19
COL2A1	12q13.11	33
SPRY2	13q31.1	3

Continuation of Table 2.1.

Gene/Locus	Chromosomal localization	Number of genotyped SNPs
BMP4	14q22-q23	4
TGFB3	14q24	8
JAG2	14q32	11
MMP25	16p13.3	7
MMP2	16q13-q21	21
CDH1	16q22.1	14
RARA	17q21	5
WNT3	17q21	17
WNT9B	17q21	12
TIMP2	17q25	26
'OFC11'	18q21**	27
BCL3	19q13.1-q13.2	4
PVRL2	19q13.2	13
CLPTM1	19q13.2-q13.3***	8
BMP2	20p12	25
MMP9	20q11.2-q13.1	6
TIMP3	22q12.3****	38
TBX22	Xq21.1	7
TIMP1	Xp11.3-p11.23	6

^ SNP - single nucleotide polymorphism;

* includes *NUDT6* gene;
** includes *SMAD2* and *SMAD4* genes;
*** includes *APOC2* gene;
**** includes *SYN3* gene

2.2.2.1.2. SNP genotyping

SNP genotyping was performed according to the standard protocol of APEX-2 genotyping method, developed in the Univerisity of Tartu, Estonia (Krjutskov et al., 2008).

Step I (DNA denaturation)

1. Mix 3 µl DNA (100 ng/µl) with 1 µl denaturation solution (5 mM Tris-HCl + 0.1 mM EDTA) and heat at 95°C for 3 minutes without lid.

Step II (I phase PCR with specific primers, volume 15 µl)

1. Prepare Master Mix (see Table 2.2).

Table 2.2.

I phase PCR Master Mix preparation

Reagent	Start concentration	Final concentration	Volume for 1 reaction, µl
ddH_2O			2.2
Buffer TrueStart original (*Fermentas*, Lithuania)	10 x	3 x	4.5
$MgCl_2$	100 mM	5.75 mM	0.86
DMSO	100 %	2 %	0.3
dNTP mix	25 mM	0.35 mM	0.2
Primer mix (650 plex)	100 nM	30 nM	6.6
TrueStart Taq (*Fermentas*, Lithuania)	5U/ µl	2 U	0.4

2. Mix all reagents, add DNA and mix gently.
3. Start the PCR using *Biometra TProfessional* (Germany) thermocycler. For temperature cycles see Table 2.3.

Table 2.3.

I phase PCR conditions

Temperature	Time	Cycles
98°C	45 seconds	1x
95°C	15 seconds	27x
62°C	30 seconds	27x
64°C	1 minute 30 seconds	27x
65°C	1 minute 30 seconds	27x
66°C	30 seconds	27x
72°C	20 seconds	27x
72°C	1 minute	1x

Step III (II phase PCR with universal primers, volume 135 µl, concentrations calculated for 150 µl)

1. Prepare Master Mix (see Table 2.4).

Table 2.4.

II phase PCR Master Mix preparation

Reagent	Volume for 1 reaction, µl
ddH_2O	50
10x B buffer (*Solis Biodyne*, Estonia)	22.5
2.5 mM dNTP mix (*Fermentas*, Lithuania)	22.5
500 pmol/ µl Primer mix (*Metabion Int. AG*, Germany)	18
25 mM $MgCl_2$ (*Fermentas*, Lithuania)	18
5U/ µl Hot Fire Pol (*Solis Biodyne*, Estonia)	4

2. Mix all reagents.
3. Add I phase PCR product and mix gently.
4. Start the PCR using *Biometra TProfessional* (Germany) thermocycler. For temperature cycles see Table 2.5.

Table 2.5.

II phase PCR conditions

Temperature	Time	Cycles
95°C	15 minutes	1x
95°C	20 seconds	30x
54°C	20 seconds	30x
72°C	5 seconds	30x
72°C	5 minutes	1x

Step IV (PCR product purification)

Purification was done using MN NucleoSpin Gel and PCR clean-up kit (*Macherey-Nagel GmbH & Co*, Germany).

1. Mix 150 µl of II phase PCR product with 500 µl binding buffer (NTI buffer), place the column into a collection tube and load the sample.

2. Centrifuge using centrifuge MiniSpin Plus (*Eppendorf*, Germany) for 30 seconds at 11 000 rpm/min, discard flow-through and place the column back into the collection tube.
3. Wash twice with 600 µl wash buffer (NT3 buffer).
4. Centrifuge for 30 seconds at 11 000 rpm/min, discard flow-through and place the column back into the collection tube.
5. Centrifuge for 1 minute at 11 000 rpm/min to remove NT3 buffer completely.
6. Place the column into a new 1.5 ml tube and add 35 µl of elution buffer (NE buffer).
7. Incubate for 1 minute at room temperature.
8. Centrifuge for 1 minute at 11 000 rpm/min.

Step V (Sample denaturation)

1. Add 4 µl ThermoSequenase (*GE Heatlhcare*, USA) 10x reaction buffer.
2. Add 0.25 µl SAP enzyme (*Fermentas*, Lithuania).
3. Incubate 15 minutes at 37°C + 6 minutes at 95°C.

Step VI

1. Prepare APEX mix (see Table 2.6).

Table 2.6.

APEX mix preparation

Reagent	Volume for 1 reaction, µl
100 µM ddNTP	0.5 each
32U/ µl ThermoSequenase (*GE Healthcare*, USA)	0.25
Dilution buffer (*GE Healthcare*, USA)	0.75

2. Mix denatured sample with 3 µl prepared APEX mix.
3. Vortex briefly and load directly 40 µl to array.
4. Hybridize array for 20 minutes at 60°C.
5. Wash slide with boiled dH_2O.
6. Image processing using *Genorama*™ 4.2.9. software (*Asper Biotech*, Estonia).

2.2.2.2. Genotyping using TaqMan chemistry

2.2.2.2.1. SNP selection

Three markers in *BMP4*, six markers in 19q13 locus and seven markers in *IRF6* gene were selected based on recent publications regarding confirmed linkage studies and associations with non-syndromic CL/CLP/CP. Detailed information about selected markers is shown in Table 2.7.

Table 2.7.

Selected SNPs for genotyping with TaqMan probes used in the study

SNP^	Chromosomal localization	Gene	SNP localization in gene	Allele*	MAF**
rs1957860	14: 53499105	*BMP4*	~6 kb downstream of gene	C/T	0.383
rs17563	14: 53487272	*BMP4*	Exon 5	A/G	0.373
rs2071047	14: 53488161	*BMP4*	Intron 4	G/A	0.406
rs35385129	19: 49854029	*PVR*	Exon 6	C/A	0.164
rs10421283	19: 49881333	*PVR/ BCL3*	~20 kb downstream of *PVR* gene and ~62 kb upstream of *BCL3* gene	G/A	0.476
rs2927438	19: 49933947	*BCL3*	~10 kb upstream of gene	G/A	0.190
rs419010	19: 50060160	*PVRL2*	Intron 1	T/C	0.484
rs2075620	19: 50171877	*CLPTM1*	Intron 6	A/G	0.362
rs875255	19: 50185475	*CLPTM1*	Intron 11	G/C	0.428
rs4844880	1: 207937539	*IRF6*	~88 kb upstream of gene	T/A	0.355
rs2013162	1: 208035307	*IRF6*	Exon 4	C/A	0.403
rs861019	1: 208042009	*IRF6*	Intron 1	A/G	0.399
rs2073487	1: 208043269	*IRF6*	Intron 1	T/C	0.402
rs642961	1: 208055893	*IRF6*	~10 kb downstream of gene	G/A	0.175
rs658860	1: 208057172	*IRF6*	~11 kb downstream of gene	C/T	0.181
rs2235371	1: 208030703	*IRF6*	Exon 6	C/T	0.138

^ SNP - single nucleotide polymorphism; * Major allele is listed first

** Minor allele frequency from http://www.ncbi.nlm.nih.gov

2.2.2.2.2. SNP genotping

Genotyping was performed using TaqMan standard assays (*Applied Biosystems*, USA) on automatic sequence-detection instruments 7500 Real-Time PCR System and ViiA™ 7 Real-Time PCR System (*Applied Biosystems*, USA).

Reactions were carried out with the use of standard conditions as suggested by the manufacturer (see Table 2.8.).

1. Prepare Master Mix (volume 10 µl) (see Table 2.8.)

Table 2.8.

Real Time PCR Master mix preparation

Reagent	Volume for 1 reaction, µl
ddH_2O	4.75
Buffer *(Applied Biosystems*, USA)	5.0
TaqMan probe *(Applied Biosystems*, USA)	0.25

2. Add 28 ng of DNA and mix all reagents.
3. Start PCR using following program (see Table 2.9.)

Table 2.9.

Real Time PCR universal program

Temperature	Time	Cycles
95°C	15 minutes	1x
95°C	15 seconds	40x
60°C	1 minute	40x
4°C	5 minutes	Infinity

2.2.2.3. MALDI-TOF (Matrix-assisted laser desorption/ionization-time-of-flight) technology

2.2.2.3.1. SNP selection

Eight SNPs (Figure 2.1.) were selected to cover the entire *BCL3* gene with 1 kb distance, taking into consideration published allele frequencies (http://www.ncbi.nlm.nih.gov/sites/entrez).

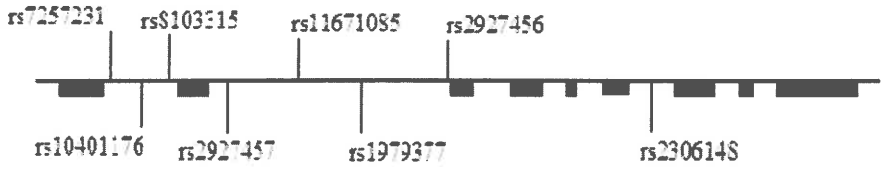

Figure 2.1. *BCL3* gene structure and approximate location of the SNPs selected in this study. Vertical lines indicate the approximate localization of selected SNPs within the *BCL3* gene. Black boxes indicate exons. Lines connecting boxes indicate introns.

2.2.2.3.2. SNP genotyping

Genotyping was performed with the use of MALDI-TOF technology using Bruker Daltonics genostrep 96 kit 10x96 (*Bruker Daltonics*, Germany) with slight modifications.

Details of selected SNPs are presented in Table 2.10.

Table 2.10.

Selected SNPs in *BCL3* gene

SNP^	Chromosomal localization	SNP localization in gene	Allele*	MAF**
rs7257231	19: 49944279	Intron 1	T/A	0.276
rs10401176	19: 49945331	Intron1	G/A	0.160
rs8103315	19: 49946008	Intron1	G/T	0.054
rs2927457	19: 49948787	Intron 2	T/G	0.051
rs11671085	19: 49949647	Intron 2	C/T	Not available
rs1979377	19: 49950842	Intron 2	G/T	0.187
rs2927456	19: 49952054	Intron 3	C/T	0.130
rs2306148	19: 49953271	Intron 6	C/T	Not available

^ SNP - single nucleotide polymorphisms; * Major allele listed first

** Minor allele frequency from http://www.ncbi.nlm.nih.gov

Details of *BCL3* gene selected SNPs, primer sequences and PCR fragments for MALDI-TOF reactions are presented in Appendix 2.

Step I (PCR)

1. Prepare Master Mix (see Table 2.11.).

Table 2.11.

PCR Master Mix for *BCL3* gene analysis

Reagent	Start concentration	Final concentration	Volume for 1 reaction, µl
ddH_2O	-	-	8.568
Buffer BD (*Solis BioDyne*, Estonia)	10x	1x	1.2
$MgCl_2$ (*Fermentas*)	25 mM	2.5 mM	1.2
dNTP mix (*Fermentas*, Lithuania)	10 mM	0.2 mM	0.24
Forward primer (*Metabion International AG*, Germany)	100 pmol/µl	0.3 pmol/µl	0.036
Reverse primer (*Metabion International AG*, Germany)	100 pmol/µl	0.3 pmol/µl	0.036
Hot FIREPol DNA polymerase (*Solis BioDyne*, Estonia)	5 U/µl	0.1 U/µl	0.24

2. Add 1 µl DNA (25 ng/µl), mix all reagents by vortex and spin shortly.
3. Start the PCR using *GeneAmp PCR System 9700* thermocycler (*Applied Biosystems*, USA). For temperature cycles see Appendix 3.
4. PCR products were checked in 1.5% agarose gel after electrophoresis at 110V for 10-20 minutes.

Step II (PCR product purification)

1. Prepare Master Mix (see Table 2.12.).

Table 2.12.

Master Mix for PCR product purification

Reagent	Volume for 96 reactions, µl
SAP enzyme (*Fermentas*, Lithuania)	29
ddH_2O	305

2. Mix all reagents by vortex and spin shortly.
3. Transfer 3 µl of reaction mix to new tube and start the PCR using *GeneAmp PCR System 9700* thermocycler (*Applied Biosystems*, USA). For temperature cycles see Table 2.13.

Table 2.13.

Temperature cycles for PCR product purification

Temperature	Time	Cycle
37°C	60 minutes	1x
94°C	20 minutes	1x

Step III (Minisequencing)

Details of minisequencing primer sequences and ddNTPs used for mini-sequencing reactions are presented in Table 2.14.

Table 2.14.

Minisequencing primer sequences and ddNTPs used for the *BCL3* genetic analysis

SNP^	Minisequencing (MS) primers	ddNTPs
rs7257231	**B**-GATCTCATATCTAT**L**TCCTTGG	ddATP; ddTTP
rs10401176	**B**-ACCAACCCATC**L**CACAGAC	ddGTP; ddATP
rs8103315	**B**-ACCTAGCAGGGGAL**C**CCAG	ddTTP; ddGTP
rs1979377	**B**-CTAACTTTTGT**L**TTTTTAGTAGAGACA	ddGTP; ddTTP
rs2927456	**B**-CTCTCTAGTCCTGC**L**TCCC	ddCTP; ddTTP

^ SNP - single nucleotide polymorphism; B - biotin; L - photolinker

1. Prepare Master Mix (see Table 2.15.).

Table 2.15.

Master mix for minisequencing reaction for *BCL3* gene analysis

Reagent	Start concentration	Final concentration	Volume for 1 reaction, µl
Buffer C (*Solis BioDyne*, Estonia)	10x	0.5x	0.5
$MgCl_2$ (*Fermentas*, Lithuania)	100 mM	1.25 mM	0.125
Primer (*BioTez Berlin-Buch GmbH*, Germany)	100 pmol/µl	1 pmol/µl	0.1
ddNTP 1 (*Fermentas*, Lithuania)	20 pmol/µl	0.2 pmol/µl	0.1
ddNTP 2 (*Fermentas*, Lithuania)	20 pmol/µl	0.2 pmol/µl	0.1
TERMI Pol DNA polymerase (*Solis BioDyne*, Estonia)	5 U/µl	0.2 U/µl	0.4
ddH_2O	-	-	3.675

2. Mix all reagents by vortex and spin shortly.
3. Start the minisequencing reaction using *GeneAmp PCR System 9700* thermocycler (*Applied Biosystems*, USA). For temperature cycles see Table 2.16.

Table 2.16.

Minisequencing program used for *BCL3* gene analysis

Temperature	Time	Cycles
95°C	2 minutes	1x
95°C	10 seconds	100x
55°C	10 seconds	100x
72°C	10 seconds	100x
72°C	5 minutes	1x

Step IV (Minisequencing product purification)

Purification was performed in 96-well plate using DNA purification kit (*Bruker Daltonics*, Germany) with slight modifications.

1. Add 4 µl binding buffer (BB) and 10 µl ddH_2O to each minisequencing product and mix 5 times by pipetting.
2. Transfer 16 µl of reaction mix to streptavidin-coated 96-well plate and incubate 30 minutes at room temperature.
3. Discard all solution and start washing process automatically with washing buffer 1 (WB1) and washing buffer 2 (WB2) ~ 1.5 hours.
4. Add 20 µl of elution buffer (EB) and insert the plate into UV incubator UV-unit CL-366™ (*Bruker Daltonics*, Germany).
5. Prepare Matrix mix (see Table 2.17.).

Table 2.17.

Matrix preparation for sample reading in MALDI-TOF mass spectrometer

Reagent	Amount
3-hydroxypicolinic acid	10 mg
DAC (dyaminoacetate)	100 µl
ddH_2O	900 µl

6. Mix all reagents and spot 1 µl of prepared matrix to 384-well iron plate (each sample is in 4 copies) and let to dry.
7. Transfer 1 µl of each sample on each dry matrix spot and let to dry.
8. Plate with dried samples is ready for analyzing in MALDI-TOF mass spectrometer (*Bruker Daltonics*, Germany).

Step V (Sample reading and data analyzing)

1. Prepare plate for calibration entering minisequencing primers sequences and calculating mass of primers and nucleotides for detecting (see Table 2.18.).

Table 2.18.

Mass of primers and nucleotides

SNP^	Primer mass, Da	Nucleotide A mass, Da	Nucleotide T mass, Da	Nucleotide C mass, Da	Nucleotide G mass, Da
rs7257231	2167.37	2464.58	2455.57	-	-
rs10401176	2154.39	2451.60	-	-	2467.60
rs8103315	1527.98	-	1816.17	-	1841.18
rs1979377	4686.02	-	4974.21	-	4999.22
rs2927456	1189.76	-	1477.95	1462.94	-

^ SNP - single nucleotide polymorphism

2. Insert the plate into MALDI-TOF mass spectrometer (*Bruker Daltonics*, Germany) and analyze all samples automatically after calibration.
3. Sample reading and data analysis is done with program "genotools^{TM} 2.0" (*Bruker Daltonics*, Germany).

3.2.3. Statistical analysis

All analyzed markers were tested for Hardy-Weinberg equilibrium in controls and affected individuals using a Pearson's chi-square test with one degree of freedom. Allele frequency differences between cleft patients and control subjects were compared for each marker using a standard chi-test with one degree of freedom. Allelic odds ratios (ORs) and 95% confidence intervals (CIs) were estimated using the standard χ^2 test, assuming a multiplicative model. The level of statistical significance was set at α=0.05 for nominal association. Haplotype-phenotype association tests were performed with the standard chi-square test using sliding windows and LD blocks approach. The PLINK software (Purcell et al., 2007) was used to perform case-control comparisons and to test for transmission distortions in the triad families. Bonferroni correction was applied for multiple testing correction considering the number of tests and variables (0.05/number of independent tests).

For performing data statistical analysis, non-syndromic CL/CLP/CP patients were divided into 2 groups - non-syndromic cleft palate (CP) patient group and non-syndromic cleft lip and cleft lip with cleft palate patient group (CL/CLP), because it has been considered that cleft lip and cleft lip with cleft palate develop at similar embryonic stages.

3. RESULTS

In this chapter results are shown according to the methods used for genotyping. All together three different genotyping methods were used.

3.1. Genotyping using APEX-2 method

Six hundred and fifty one markers in 44 genes were analyzed for 106 cleft lip and cleft lip with cleft palate (CL/CLP) patients, 29 cleft palate (CP) patients and 182 healthy and unrelated individuals as controls. After data quality cleaning nine markers were excluded based on HWE test (p value <0.05), twenty-one markers failed missingness test (>95%) in CL/CLP group and nine markers in CP group, and 1 control sample was removed because of individual missingness treshold of <10%. The overall genotype rate was ~ 99%.

In the Table 3.1. all markers which remained associated after correction for multiple testing in the CL/CLP and CP sample set are presented. The best results of single marker association analysis (P≤0.05) are shown in Appendix 4.

3.1.1. CL/CLP group

The strongest association with CL/CLP was found for SNP rs16824948, which is located in *SKI* gene, where the allele T was associated with increased risk (p-value = $0.0013x10^{-14}$; OR = 6.376; 95% CI = 4.039-10.07). Obtained association remained statistically significant after Bonferroni correction.

SNP rs11655598 in *WNT3* gene showed very strong association (p-value = $0.0053x10^{-11}$; OR = 5.925; 95% CI = 3.593-9.772) with CL/CLP, which remained significant after correction for multiple testing. Allele G was associated with increased risk for CL/CLP.

Also allele C for SNP rs7829058 in *FGFR1* gene showed increased risk (p-value = $0.0024x10^{-5}$; OR = 7.991; 95% CI = 3.435-18.59) for non-syndromic CL/CLP.

3.1.2. CP group

The strongest association with CP was found for SNP rs11655598, located in *WNT3* gene, where the allele G was associated with increased risk (p-value = $0.0039x10^{-11}$; OR = 9.495; 95% CI = 4.879-18.34).

SNP rs16824948 in *SKI* gene showed very strong association (p-value = $0.0011x10^{-7}$; OR = 6.777; 95% CI = 3.577-12.84) with CP, which remained significant after correction for multiple testing. Allele T was associated with increased risk for CP.

Rs7829058 in *FGFR1* gene showed increased risk (p-value = $0.0002x10^{-6}$; OR = 13.16; 95% CI = 4.93-35.1) for non-syndromic cleft palate, whereas the allele C was associated with increased risk for cleft palate phenotype.

Table 3.1.

Most significant results of single-marker association analysis associated with non-syndromic CL/CLP and CP

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
CL/CLP									
1	*SKI*	rs16824948	2176080	C/T	0.382	0.088	$0.0013x10^{-14}$	6.376	4.039-10.07
8	*FGFR1*	rs7829058	38451252	G/C	0.137	0.019	$0.0024x10^{-5}$	7.991	3.435-18.59
17	*WNT3*	rs11655598	42223260	C/G	0.307	0.069	$0.0053x10^{-11}$	5.925	3.593-9.772
CP									
1	*SKI*	rs16824948	2176080	C/T	0.397	0.088	$0.0011x10^{-7}$	6.777	3.577-12.84
8	*FGFR1*	rs7829058	38451252	G/C	0.207	0.019	$0.0002x10^{-6}$	13.16	4.93-35.1
17	*WNT3*	rs11655598	42223260	C/G	0.414	0.069	$0.0039x10^{-11}$	9.459	4.879-18.34

* Chr - chromosome;
^ SNP - single nucleotide polymorphism;
# Major allele is listed first;
** MAF - minor allele frequency;
^^ OR - odds ratio;
95% CI - 95% confidence interval

Haplotype-association analysis was performed using two different approaches. Haplotype analysis was applied using two to five SNP slinding window approach for the genes *SKI*, *WNT3* and *FGFR1* which were strongly associated with CL/CLP and CP phenotype in single-marker association analysis. Second approach was haplotype-based association analysis within LD blocks for all genes.

In the Table 3.2. and Table 3.3. best results (p value ≤ 0.0001) of haplotype analysis using sliding window in *SKI* gene are presented.

The strongest association with CL/CLP in *SKI* gene was found for rs16824948-rs903910 (TC) (p value = 0.0392×10^{-15}), rs16824948-rs903910-rs4648625 (TCT) (p value = 0.0062×10^{-14}), rs263533-rs16824948-rs903910 (CTC) (p value = 0.0034×10^{-8}), rs263533-rs16824948-rs903910-rs4648625 (CTCT) (p value = 0.0132×10^{-8}), rs262683-rs2460000-rs263533-rs16824948-rs903910 (TGTTC) (p value = 0.0172×10^{-8}), rs263533-rs16824948 (CT) (p value = 0.0176×10^{-8}), rs16824948-rs903910-rs4648625-rs6673129-rs12045693 (TCTCA) (p value = 0.0282×10^{-8}), rs16824948-rs903910-rs4648625-rs6673129 (TCTC) (p value = 0.0092×10^{-7}) and rs2460000-rs263533-rs16824948-rs903910-rs4648625 (GTTCT) (p value = 0.0106×10^{-7}) haplotypes, which were associated with higher risk of disease. Very strong association with CP was found for rs2460000-rs263533-rs16824948-rs260507-rs903910 (GTTCC) (p value = 0.0029×10^{-16}), rs16824948-rs260507-rs903910-rs4648625-rs6673129 (TCCTC) (p value = 0.0045×10^{-12}), rs16824948-rs260507-rs903910-rs4648625 (TCCT) (p value = 0.0075×10^{-12}), rs16824948-rs260507-rs903910 (TCC) (p value = 0.0126×10^{-12}), rs263533-rs16824948-rs260507-rs903910-rs4648625 (TTCCT), (p value = 0.0034×10^{-11}), rs263533-rs16824948-rs260507-rs903910 (TTCC), (p value = 0.0061×10^{-11}) and rs2460000-rs263533-rs16824948-rs260507 (GTTC) (p value = 0.0136×10^{-11}) which all were associated with higher risk for CP.

Table 3.2.

The best results of haplotype analysis in *SKI* gene associated with CL/CLP

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs6665593	rs262683	rs2460000	rs263533	rs16824948	*	*	*
WIN1	G	C	A	C	T	0.11	0.007	$0.0163x10^{-5}$
WIN1	A	T	G	T	T	0.095	0.006	$0.0105x10^{-5}$
*	rs262683	rs2460000	rs263533	rs16824948	rs903910	*	*	*
WIN2	T	G	T	T	C	0.112	0.001	$0.0172x10^{-8}$
WIN2	C	G	C	T	C	0.09	0.003	$0.0044x10^{-5}$
WIN2	C	A	C	T	T	0.072	0.007	$0.0126x10^{-3}$
WIN2	C	A	C	T	C	0.044	0.0007	$0.0596x10^{-3}$
WIN2	T	G	T	T	*	0.113	0.006	$0.0036x10^{-5}$
WIN2	C	A	C	T	*	0.11	0.011	$0.0067x10^{-4}$
*	rs2460000	rs263533	rs16824948	rs903910	rs4648625	*	*	*
WIN3	G	T	T	C	T	0.109	0.003	$0.0106x10^{-7}$
WIN3	G	C	T	C	T	0.084	0.002	$0.0111x10^{-5}$
WIN3	A	C	T	T	T	0.07	0.005	$0.0093x10^{-3}$
WIN3	A	C	T	C	T	0.051	0.0009	$0.0137x10^{-3}$
WIN3	G	T	T	C	*	0.105	0.003	$0.0188x10^{-6}$
WIN3	G	C	T	C	*	0.088	0.003	$0.0052x10^{-4}$

Continuation of Table 3.2.

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN3	A	C	T	T	*	0.083	0.006	$0.0082x10^{-3}$
WIN3	A	C	T	C	*	0.054	0.0009	$0.0078x10^{-2}$
WIN3	A	C	T	*	*	0.133	0.01	$0.0056x10^{-6}$
WIN3	G	T	T	*	*	0.102	0.007	$0.0042x10^{-4}$
*	rs263533	rs16824948	rs903910	rs4648625	rs6673129	*	*	*
WIN4	T	T	C	T	C	0.095	0.007	$0.0148x10^{-4}$
WIN4	C	T	C	T	T	0.079	0.003	$0.0042x10^{-4}$
WIN4	C	T	C	T	C	0.047	0	$0.0357x10^{-3}$
WIN4	C	T	C	T		0.13	0.003	$0.0132x10^{-8}$
WIN4	T	C	T	T	*	0.095	0.007	$0.0164x10^{-4}$
WIN4	C	T	C	*	*	0.137	0.003	$0.0034x10^{-8}$
WIN4	T	T	C	*	*	0.092	0.008	$0.0355x10^{-4}$
WIN4	C	T	*	*	*	0.291	0.079	$0.0176x10^{-8}$
WIN4	T	T	*	*	*	0.091	0.009	$0.0115x10^{-3}$
*	rs16824948	rs903910	rs4648625	rs6673129	rs12045693	*	*	*
WIN5	T	C	T	C	A	0.116	0.003	$0.0282x10^{-8}$
WIN5	T	C	T	T	C	0.087	0.002	$0.0043x10^{-5}$

Continuation of Table 3.2.

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN5	T	T	T	C	C	0.071	0.011	0.0001
WIN5	T	C	T	C	*	0.134	0.007	0.0092×10^{-7}
WIN5	T	C	T	T	*	0.092	0.005	0.0012×10^{-3}
WIN5	T	C	T	*	*	0.228	0.013	0.0062×10^{-14}
WIN5	C	C	T	*	*	0.476	0.65	0.0475×10^{-2}
WIN5	T	C	*	*	*	0.233	0.014	0.0392×10^{-15}
WIN5	C	C	*	*	*	0.465	0.648	0.0171×10^{-2}

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4; WIN5 - sliding window 5
^ SNP - single nucleotide polymorphism; * Empty cell

Table 3.3.

The best results of haplotype analysis in *SKI* gene associated with CP

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs6665593	rs262683	rs2460000	rs263533	rs16824948	*	*	*
WIN1	G	T	G	T	T	0.112	0.001	$0.0134x10^{-7}$
WIN1	G	C	A	C	T	0.098	0.004	0.00005
*	rs262683	rs2460000	rs263533	rs16824948	rs260507	*	*	*
WIN2	T	G	T	T	C	0.167	0.004	$0.0048x10^{-9}$
WIN2	C	A	C	T	G	0.107	0.007	$0.0066x10^{-3}$
WIN2	T	G	T	T	*	0.153	0.005	$0.0271x10^{-8}$
WIN2	C	A	C	T	*	0.112	0.008	$0.0092x10^{-3}$
*	rs2460000	rs263533	rs16824948	rs260507	rs903910	*	*	*
WIN3	G	T	T	C	C	0.261	0.004	$0.0029x10^{-16}$
WIN3	G	T	T	C	*	0.203	0.007	$0.0136x10^{-11}$
WIN3	A	C	T	G	*	0.095	0.004	$0.0066x10^{-3}$
WIN3	G	T	T	*	*	0.187	0.008	$0.0099x10^{-9}$
WIN3	A	C	T	*	*	0.102	0.006	$0.0006x10^{-2}$
*	rs263533	rs16824948	rs260507	rs903910	rs4648625	*	*	*
WIN4	T	T	C	C	T	0.218	0.008	$0.0034x10^{-11}$
WIN4	T	T	C	C	*	0.219	0.008	$0.0061x10^{-11}$

Continuation of Table 3.3.

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN4	T	T	C	*	*	0.191	0.012	$0.0113x10^{-8}$
WIN4	T	T	*	*	*	0.185	0.013	$0.0135x10^{-7}$
*	rs16824948	rs260507	rs903910	rs4648625	rs6673129	*	*	*
WIN5	T	C	C	T	C	0.219	0.006	$0.0045x10^{-12}$
WIN5	T	C	C	T	*	0.218	0.006	$0.0075x10^{-12}$
WIN5	T	C	C	*	*	0.22	0.007	$0.0126x10^{-12}$
WIN5	T	C	*	*	*	0.207	0.01	$0.0059x10^{-9}$

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - slidind window 3; WIN4 - sliding window 4; WIN5 - sliding window 5
^ SNP - single nucleotide polymorphism; * Empty cell

Table 3.4. presents the best haplotype-based association results (p value $\leq$0.0001) in *FGFR1* gene.

We found strongest association between *FGFR1* haplotypes rs7829058-rs13279569 (CG) (p value = 0.0045×10^{-5}), rs6996321-rs7829058-rs13279569 (GCG) (p value = 0.0185×10^{-5}) and CL/CLP. Both haplotypes were associated with higher risk for CL/CLP. Strongest association was found with CP for three *FGFR1* haplotypes - rs7829058-rs13279569 (CG) (p value = 0.0331×10^{-7}), rs6996321-rs7829058 (GC) (p value = 0.0021×10^{-4}) and rs6996321-rs7829058-rs13279569-rs328300 (GCGT) (p value = 0.0026×10^{-4}), which were associated with higher risk too.

Table 3.4.

The best results of haplotype analysis in *FGFR1* gene associated with non-syndromic CL/CLP and CP

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
CL/CLP								
*	rs6987534	rs6474354	rs7012413	rs6996321	rs7829058	*	*	*
WIN6	C	T	C	G	C	0.076	0.007	$0.0083x10^{-2}$
	rs6474354	rs7012413	rs6996321	rs7829058	rs13279569	*	*	*
WIN7	T	C	G	C	G	0.072	0.006	$0.0106x10^{-3}$
WIN7	T	C	G	C	*	0.072	0.006	$0.0096x10^{-2}$
*	rs7012413	rs6996321	rs7829058	rs13279569	rs328300	*	*	*
WIN8	C	G	C	G	G	0.097	0.018	$0.0254x10^{-3}$
WIN8	C	G	G	G	G	0.002	0.079	0.0001
WIN8	C	G	C	G	*	0.124	0.019	$0.0263x10^{-4}$
WIN8	C	G	C	*	*	0.124	0.019	$0.0022x10^{-3}$
*	rs6996321	rs7829058	rs13279569	*	*	*	*	*
WIN9	G	C	G	*	*	0.138	0.019	$0.0185x10^{-5}$
WIN9	G	C	*	*	*	0.126	0.019	$0.0158x10^{-4}$
*	rs7829058	rs13279569	rs328300	*	*	*	*	*
WIN10	C	G	2	*	*	0.052	0.003	0.0001
WIN10	C	G	*	*	*	0.147	0.019	$0.0045x10^{-5}$

Continuation of Table 3.4.

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
CP								
*	rs7012413	rs6996321	rs7829058	rs13279569	rs328300	*	*	*
WIN8	C	G	C	G	T	0.079	0.003	$0.0037x10^{-2}$
WIN8	C	G	C	G	*	0.137	0.018	$0.0085x10^{-2}$
WIN8	C	G	C	*	*	0.137	0.018	$0.0073x10^{-2}$
*	rs6996321	rs7829058	rs13279569	rs328300	*	*	*	*
WIN9	G	C	G	T	*	0.106	0.003	$0.0026x10^{-4}$
WIN9	G	C	G	*	*	0.166	0.018	$0.0118x10^{-4}$
WIN9	G	C	*	*	*	0.165	0.018	$0.0021x10^{-4}$
*	rs7829058	rs13279569	rs328300	*	*	*	*	*
WIN10	C	G	T	*	*	0.095	0.004	$0.0062x10^{-3}$
WIN10	C	G	*	*	*	0.185	0.019	$0.0331x10^{-7}$

WIN6 - sliding window 6; WIN7 - sliding window 7; WIN8 - sliding window 8; WIN9 - sliding window 9; WIN10 - sliding window 10
^ SNP - single nucleotide polymorphism; * Empty cell

In the Table 3.5. best results of haplotype analysis of *WNT3* gene are presented.

Strongest association between CL/CLP and *WNT3* gene haplotypes were found for rs199496-rs11658976-rs11655598-rs12452064-rs199494 (GAGGA) (p value = $0.0248x10^{-10}$), rs11655598-rs12452064-rs199494-rs7218567-rs111769 (GGACT) (p value = $0.0034x10^{-9}$), rs11658976-rs11655598-rs12452064-rs199494-rs7218567 (AGGAC) (p value = $0.0084x10^{-9}$) and rs11655598-rs12452064 (GG) (p value = $0.0098x10^{-9}$), which were associated with increased risk for this cleft phenotype. Strongest association with CP and *WNT3* gene was found for haplotypes rs11655598-rs12452064 (GG) (p value = $0.0049x10^{-10}$), rs11655598-rs12452064-rs199494-rs7218567-rs111769 (GGACT) (p value = $0.0006x10^{-9}$), rs11655598-rs12452064-rs199494-rs7218567 (GGAC) (p value = $0.0366x10^{-10}$), rs11655598-rs12452064-rs199494 (GGA) (p value = $0.0041x10^{-9}$) and rs11658976-rs11655598 (AG) (p value = $0.0078x10^{-9}$). All haplotypes were associated with higher risk for CP.

Table 3.5.

The best results of haplotype analysis in *WNT3* gene associated with non-syndromic CL/CLP and CP

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
CL/CLP								
*	rs916888	rs199497	rs199496	rs11658976	rs11655598	*	*	*
WIN2	T	T	G	A	G	0.255	0.067	$0.0293x10^{-8}$
	rs199497	rs199496	rs11658976	rs11655598	rs12452064			
WIN3	T	G	A	G	G	0.266	0.064	$0.0018x10^{-8}$
WIN3	T	G	A	G	*	0.26	0.066	$0.0067x10^{-7}$
*	rs199496	rs11658976	rs11655598	rs12452064	rs199494	*	*	*
WIN4	G	A	G	G	A	0.277	0.063	$0.0248x10^{-10}$
WIN4	G	A	C	G	A	0.127	0.29	$0.0115x10^{-3}$
WIN4	G	A	G	G	*	0.265	0.062	$0.0111x10^{-8}$
WIN4	G	A	C	G	*	0.136	0.297	$0.0145x10^{-2}$
WIN4	G	A	G	*	*	0.264	0.065	$0.0024x10^{-7}$
*	rs11658976	rs11655598	rs12452064	rs199494	rs7218567	*	*	*
WIN5	A	G	G	A	C	0.274	0.064	$0.0084x10^{-9}$
WIN5	A	C	G	A	C	0.146	0.31	$0.0022x10^{-2}$
WIN5	A	G	G	A	*	0.266	0.06	$0.0004x10^{-7}$
WIN5	A	C	G	A	*	0.14	0.304	$0.0118x10^{-2}$

Continuation of Table 3.5.

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN5	A	G	G	*	*	0.259	0.06	0.0159×10^{-8}
WIN5	A	C	G	*	*	0.229	0.407	0.0162×10^{-2}
WIN5	A	G	*	*	*	0.259	0.064	0.0393×10^{-8}
*	rs11655598	rs12452064	rs199494	rs7218567	rs111769	*	*	*
WIN6	G	G	A	C	T	0.29	0.07	0.0034×10^{-9}
WIN6	C	G	A	C	T	0.144	0.297	0.0057×10^{-2}
WIN6	G	G	A	C	*	0.285	0.069	0.0311×10^{-9}
WIN6	C	G	A	C	*	0.146	0.326	0.0293×10^{-3}
WIN6	G	G	A	*	*	0.274	0.067	0.0107×10^{-8}
WIN6	C	G	A	*	*	0.144	0.329	0.0135×10^{-3}
WIN6	G	G	*	*	*	0.289	0.068	0.0098×10^{-9}
WIN6	C	G	*	*	*	0.243	0.455	0.0054×10^{-3}
CP								
*	rs916888	rs199497	rs199496	rs11658976	rs11655598	*	*	*
WIN2	T	T	G	A	G	0.338	0.066	0.0074×10^{-6}
*	rs199497	rs199496	rs11658976	rs11655598	rs12452064	*	*	*
WIN3	T	G	A	G	G	0.357	0.063	0.0207×10^{-8}
WIN3	T	G	A	G	*	0.357	0.066	0.0387×10^{-8}

Continuation of Table 3.5.

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs199496	rs11658976	rs11655598	rs12452064	rs199494	*	*	*
WIN4	G	A	G	G	A	0.277	0.063	$0.0248x10^{-9}$
WIN4	G	A	G	G	*	0.353	0.061	$0.0151x10^{-8}$
WIN4	G	A	G	*	*	0.352	0.065	$0.0046x10^{-7}$
*	rs11658976	rs11655598	rs12452064	rs199494	rs7218567	*	*	*
WIN5	A	G	G	A	C	0.353	0.057	$0.0063x10^{-8}$
WIN5	A	G	G	A	*	0.348	0.056	$0.0089x10^{-8}$
WIN5	A	G	G	*	*	0.352	0.058	$0.0083x10^{-8}$
WIN5	A	G	*	*	*	0.383	0.064	$0.0078x10^{-9}$
*	rs11655598	rs12452064	rs199494	rs7218567	rs111769	*	*	*
WIN6	G	G	A	C	T	0.411	0.07	$0.0006x10^{-9}$
WIN6	G	G	A	C	*	0.404	0.069	$0.0366x10^{-10}$
WIN6	G	G	A	*	*	0.402	0.068	$0.0041x10^{-9}$
WIN6	G	G	*	*	*	0.416	0.069	$0.0049x10^{-10}$

WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4; WIN5 - sliding window 5; WIN6 - sliding window 6
^ SNP - single nucleotide polymorphism; * Empty cell

After performing haplotype-based association analysis within LD blocks for all genes, all together 114 different haploblocks were generated compared CL/CLP patients and controls and 111 haploblocks - compared CP patients and controls. Only one haplotype in *FGF1* gene showed strong association (p value ≤ 0.0001) with CL/CLP, but not with CP.

Haplotype rs34002-rs250092-rs34010-rs250103-rs34013 (TGGAT), which was associated with higher risk of this cleft phenotype, is shown in the Table 3.6.

Table 3.6.

Results of haplotype analysis within LD blocks in *FGF1* gene associated with CL/CLP

Haplo -type	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs-34002	rs-250092	rs-34010	rs-250103	rs-34013	*	*	*
H1	T	G	G	A	T	0.105	0.001	0.0082×10^{-7}
H2	T	G	T	A	T	0.222	0.359	0.0007

H1 - haplotype 1; H2 - haplotype 2
^ SNP - single nucleotide polymorphism; * Empty cell

3.2. Genotyping with MALDI-TOF technology

3.2.1. *BCL3* gene

Eight markers were selected for genotyping with MALDI-TOF technology and three SNPs (rs2927457, rs11671085, and rs2306148) were not informative, therefore were not considered for further analysis.

To perform case-control association study five SNPs in *BCL3* gene were analyzed for 129 non-syndromic cleft lip and cleft lip with or without cleft palate (CL/CLP) patients, 39 non-syndromic cleft palate (CP) patients and 335 unrelated unaffected individuals as control group in Latvian population. In order to analyze these markers in another population from European origin, 606 DNA samples from a Brazilian population (294 non-syndromic CL/CLP cases, 44 CP cases and 268 unrelated and unaffected individuals with no family history of CL/CLP/CP) were studied. Transmission disequilibrium test (TDT) was performed for 109 trios (affected sib with both parents), out of all 85 sibs and their parents were divided in CL/CLP group and 24 trios - in CP group. After data quality cleaning for case-control analysis 28 individuals from CL/CLP group (14 idividuals from cases and 14 individuals from controls) and 22 individuals from CP group (eight individuals from cases and 14 individuals from controls) in Latvian population were removed because of individual missingness treshold of <10%. The overall genotype rate was 100%. After performed data cleaning in Brazilian population, 20 CL/CLP patients, two CP patients and 76 individuals from controls were excluded because of individual missingness treshold of <10%. The overall genotype rate was ~90%. After data cleaning for family based association analysis (TDT test) 32 individuals from CL/CLP group (nine individuals from cases and 23 individuals from controls) and 15 individuals from CP group (four individuals from cases and 11 individuals from parents as controls) in Latvian population were removed based on individual missingness treshold of <10%. The overall genotype rate was 100%.

In the Table 3.7. results for case-control comparisons with CL/CLP and CP in Latvian population are showed, while in Table 3.8. results of single marker association analysis with CL/CLP and CP in Brazilian population are showed.

We did not find any significant association of analyzed markers in *BCL3* gene between CL/CLP or CP cases and controls in population from Latvia or Brazil. The

only suggestive evidence for association was for SNP rs10401176 with CL/CLP in the Latvian cohort (p-value = 0.042; OR = 0.609; 95% CI = 0.377-0.986) where allele A was associated with decreased risk for non-syndromic CL/CLP, but it was not significant after Bonferroni correction.

Table 3.7.

BCL3 gene results of case-control analysis associated with non-syndromic CL/CLP and CP in Latvian population

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
CL/CLP									
19	*BCL3*	rs7257231	49944279	A/T	0.165	0.181	0.5976	0.897	0.6-1.341
19	*BCL3*	rs10401176	49945331	G/A	0.1	0.154	0.042	0.609	0.377-0.986
19	*BCL3*	rs8103315	49946008	G/T	0.1	0.075	0.23	1.375	0.816-2.317
19	*BCL3*	rs1979377	49950842	T/G	0.096	0.108	0.6148	0.878	0.53-1.456
19	*BCL3*	rs2927456	49952054	C/T	0.057	0.073	0.391	0.758	0.403-1.43
CP									
19	*BCL3*	rs7257231	49944279	A/T	0.129	0.181	0.3079	0.672	0.311-1.45
19	*BCL3*	rs10401176	49945331	G/A	0.097	0.154	0.2254	0.588	0.247-1.401
19	*BCL3*	rs8103315	49946008	G/T	0.032	0.075	0.2134	0.413	0.098-1.74
19	*BCL3*	rs1979377	49950842	T/G	0.048	0.108	0.1426	0.422	0.129-1.383
19	*BCL3*	rs2927456	49952054	C/T	0.016	0.073	0.0886	0.208	0.028-1.531

* Chr - chromosome; ^ SNP - single nucleotide polymorphism; # Major allele is listed first
** MAF - minor allele frequency; ^^ OR - odds ratio; ## 95% CI - 95% confidence interval

Table 3.8.

BCL3 gene results of case-control analysis associated with non-syndromic CL/CLP and CP in Brazilian population

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
CL/CLP									
19	*BCL3*	rs7257231	49944279	A/T	0.277	0.302	0.4013	0.884	0.663-1.179
19	*BCL3*	rs10401176	49945331	G/A	0.125	0.117	0.72	1.076	0.72-1.608
19	*BCL3*	rs8103315	49946008	G/T	0.109	0.094	0.4713	1.174	0.759-1.817
19	*BCL3*	rs2927456	49952054	C/T	0.111	0.078	0.0929	1.478	0.935-2.337
CP									
19	*BCL3*	rs7257231	49944279	A/T	0.274	0.302	0.6075	0.871	0.514-1.475
19	*BCL3*	rs10401176	49945331	G/A	0.191	0.117	0.0707	1.773	0.947-3.319
19	*BCL3*	rs8103315	49946008	G/T	0.155	0.094	0.1016	1.76	0.888-3.486
19	*BCL3*	rs2927456	49952054	C/T	0.095	0.078	0.603	1.242	0.548-2.815

* Chr - chromosome;
^ SNP - single nucleotide polymorphism;
# Major allele is listed first;
** MAF - minor allele frequency;
^^ OR - odds ratio;
95% CI - 95% confidence interval

Haplotype based association analysis was performed to find any possible association with CL/CLP and CP in Latvian and Brazilian populations.

The strongest associations with CL/CLP in Latvian population were found for *BCL3* gene haplotypes rs7257231-rs10401176-rs810315-rs1979377 (AGTT) (p-value = 0.0005), rs7257231-rs10401176-rs810315 (AGT) (p-value = 0.0006), rs7257231-rs10401176-rs810315-rs1979377-rs2927456 (AGTTT) (p-value = 0.0007) and rs10401176-rs810315-rs1979377 (GTT) (p-value = 0.0009), which were associated with an increased risk of CL/CLP.

Table 3.9. shows best results of haplotype analysis (p value $\leq$0.001) with CL/CLP in *BCL3* gene. All results of performed analysis are shown in Appendix 5.

Haplotype rs7257231-rs10401176 (TA) showed weak association with CP (p-value = 0.0345) in Latvian population and it was associated with lower risk of this cleft phenotype (see Appendix 6).

Haplotype analysis in Brazilian population showed that haplotypes rs10401176-rs810315 (GG) (p-value = 0.0078), rs7257231-rs10401176-rs810315 (TAG) (p-value = 0.0321) and rs10401176-rs810315-rs2927456 (GGC) (p-value = 0.0357) revealed borderline association with CP, but haplotypes did not showed significant association with CL/CLP (see Appendix 7 and Appendix 8).

The transmission disequilibrium test was carried out in Latvian non-syndromic CL/CLP and CP individuals and their parents in order to identify transmission distortions. No association was found for any analyzed markers with CL/CLP or CP. Obtained results are presented in Table 3.10.

Table 3.9.

The best results of haplotype analysis in *BCL3* gene between CL/CLP patients and controls in Latvian population

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs7257231	rs10401176	rs8103315	rs1979377	rs2927456	*	*	*
WIN1	A	G	T	T	T	0.101	0.034	0.0007
*	rs7257231	rs10401176	rs8103315	rs1979377	*	*	*	*
WIN1	A	G	T	T	*	0.099	0.039	0.0005
*	rs7257231	rs10401176	rs8103315	*	*	*	*	*
WIN1	A	G	T	*	*	0.098	0.038	0.0006
*	rs10401176	rs8103315	rs1979377	*	*	*	*	*
WIN2	G	T	T	*	*	0.096	0.039	0.0009

WIN1 - sliding window 1; WIN2 - sliding window 2
^ SNP - single nucleotide polymorphism; * Empty cell

Table 3.10.

Transmission distortion results in *BCL3* gene for Latvian CL/CLP and CP individuals

Chr*	Gene	SNP^	Location	Alleles#	Transmitted minor allele count	Untransmitted allele count	p-value	OR^^	95% CI##
CL/CLP									
19	*BCL3*	rs7257231	49944279	A/T	17	28	0.1011	0.607	0.332-1.109
19	*BCL3*	rs10401176	49945331	G/A	12	18	0.2087	0.667	0.321-1.384
19	*BCL3*	rs8103315	49946008	G/T	13	20	0.3035	0.65	0.65-0.323
19	*BCL3*	rs1979377	49950842	T/G	9	10	0.6547	0.9	0.366-2.215
19	*BCL3*	rs2927456	49952054	C/T	5	6	0.763	0.833	0.254-2.731
CP									
19	*BCL3*	rs7257231	49944279	A/T	3	6	0.3173	0.5	0.125-1.999
19	*BCL3*	rs10401176	49945331	G/A	2	2	1	1	0.141-7.099
19	*BCL3*	rs8103315	49946008	G/T	2	2	1	1	0.141-7.099
19	*BCL3*	rs1979377	49950842	T/G	1	2	0.5637	0.5	0.045-5.514
19	*BCL3*	rs2927456	49952054	C/T	0	2	1	NAN	NAN

* Chr - chromosome; ^ SNP - single nucleotide polymorphism; # Major allele is listed first
^^ OR - odds ratio; ## 95% CI - 95% confidence interval

3.3. Genotyping using TaqMan chemistry

3.3.1. 19q13 locus

In present study were analyzed seven markers in 19q13 locus, which contains *PVR*, *BCL3*, *PVRL2* and *CMPTM1* genes. We performed case control comparisons for all seven markers between 113 non-syndromic CL/CLP/CP patients (86 patients with CL/CLP and 27 patients with CP only) and 148 unrelated and unaffected individuals as controls in Latvian population. Transmission disequilibrium test (TDT) was performed for 66 trios (affected sib with both parents), out of all 52 sibs and their parents were divided in CL/CLP group and 14 trios - in CP group. After data quality cleaning one marker was excluded based on HWE test (p value <0.05), for case-control analysis 7 individuals from CL/CLP group (1 idividual from cases and 6 individuals from controls) and 8 individuals from CP group (2 individuals from cases and 6 individuals from controls) were excluded because of individual missingnes <10%. The overall genotype rate was 100%. After data quality cleaning for family based association analysis 3 individuals from CL/CLP group (1 individual from cases and 2 individuals from parents as controls) and 2 individuals from CP group (1 individual from cases and 1 individual from parents as controls) were removed based on individual missingnes <10%. The overall genotype rate was 100%.

Table 3.11. presents results for case-control analysis with CL/CLP and CP in Latvian population.

We did not find any significant association of analyzed markers in 19q13 locus between CL/CLP or CP cases and controls.

In present study haplotype based association analysis in 19q13 locus was performed. We did not find any association of analyzed haplotypes with CL/CLP in Latvian population. Haplotypes rs419010-rs2075620 (CG) (p-value = 0.0156), rs419010-rs2075620-rs875255 (CGC) (p-value = 0.0161), rs2927438-rs419010-rs2075620-rs875255 (GCGC) (p-value = 0.0279) and rs2927438-rs419010-rs2075620 (GCG) (p-value = 0.0305) showed very weak association with CP and these haplotypes were associated with an increased risk of CP (see Appendix 9 and Appendix 10).

The transmission disequilibrium test was carried out in Latvian non-syndromic cleft lip with or without cleft palate and isolated cleft palate individuals and their

parents in order to identify transmission distortions. Borderline association was found only for one marker in *BCL3* gene (rs10421283) (p-value = 0.0477) with CL/CLP but not with CP, which did not remain significant after Bonferroni correction. Obtained results are presented in Table 3.12.

Table 3.11.

19q13 locus results of case-control analysis associated with non-syndromic CL/CLP and CP in Latvian population

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
CL/CLP									
19	*PVR*	rs35385129	49854029	C/A	0.182	0.204	0.5699	0.869	0.535-1.411
19	*BCL3*	rs10421283	49881333	G/A	0.377	0.426	0.298	0.813	0.551-1.2
19	*BCL3*	rs2927438	49933947	G/A	0.206	0.148	0.111	1.494	0.91-2.452
19	*PVRL2*	rs419010	50060160	T/C	0.371	0.419	0.3082	0.816	0.553-1.206
19	*CLPTM1*	rs2075620	50171877	A/G	0.235	0.229	0.8752	1.037	0.661-1.625
19	*CLPTM1*	rs875255	50185475	G/C	0.429	0.465	0.4635	0.867	0.591-1.271
CP									
19	*PVR*	rs35385129	49854029	C/A	0.24	0.204	0.5666	1.23	0.605-2.503
19	*BCL3*	rs10421283	49881333	G/A	0.44	0.426	0.8542	1.058	0.578-1.94
19	*BCL3*	rs2927438	49933947	G/A	0.12	0.148	0.6042	0.786	0.315-1.959
19	*PVRL2*	rs419010	50060160	T/C	0.54	0.419	0.1118	1.628	0.89-2.978
19	*CLPTM1*	rs2075620	50171877	A/G	0.34	0.229	0.0923	1.736	0.909-3.316
19	*CLPTM1*	rs875255	50185475	G/C	0.56	0.465	0.214	1.466	0.80-2.684

* Chr - chromosome; ^ SNP - single nucleotide polymorphism; # Major allele is listed first
** MAF - minor allele frequency; ^^ OR - odds ratio; ## 95% CI - 95% confidence interval

Table 3.12.

Transmission distortion results in 19q13 locus for Latvian CL/CLP and CP individuals

Chr*	Gene	SNP^	Location	Alleles#	Transmitted minor allele count	Untransmitted allele count	p-value	OR^^	95% CI##
CL/CLP									
19	*PVR*	rs1058402	49842454	G/A	6	2	0.1573	3	0.606-14.86
19	*PVR*	rs35385129	49854029	C/A	12	21	0.0864	0.571	0.281-1.161
19	*BCL3*	rs10421283	49881333	G/A	18	31	0.0477	0.581	0.325-1.038
19	*BCL3*	rs2927438	49933947	G/A	14	16	0.715	0.875	0.427-1.793
19	*PVRL2*	rs419010	50060160	T/C	22	30	0.2673	0.733	0.423-1.271
19	*CLPTM1*	rs2075620	50171877	A/G	18	21	0.631	0.857	0.457-1.609
19	*CLPTM1*	rs875255	50185475	G/C	22	28	0.3961	0.786	0.45-1.373
CP									
19	*PVR*	rs1058402	49842454	G/A	1	2	0.5637	0.5	0.045-5.514
19	*PVR*	rs35385129	49854029	C/A	3	5	0.4795	0.6	0.143-2.511
19	*BCL3*	rs10421283	49881333	G/A	2	4	0.4142	0.5	0.092-2.73
19	*BCL3*	rs2927438	49933947	G/A	4	4	1	1	0.25-3.998
19	*PVRL2*	rs419010	50060160	T/C	7	7	1	1	0.351-2.851
19	*CLPTM1*	rs2075620	50171877	A/G	8	3	0.1317	2.667	0.708-10.05

* Chr - chromosome; ^ SNP - single nucleotide polymorphism; # Major allele is listed first; ^^ OR - odds ratio; ## 95% CI - 95% confidence interval

3.3.2. *BMP4* gene

To perform case-control comparisons in *BMP4* gene three SNPs were analyzed for 127 cleft lip with or without cleft palate (CL/CLP) patients, 37 cleft palate (CP) patients and 190 unrelated and healthy individuals with no family history of non-syndromic CL/CLP/CP as controls in Latvian population. Transmission disequilibrium test was performed for 65 trios (affected sib with both parents), out of all 38 sibs and their parents were divided in CL/CLP group and 27 trios - in CP group. After data quality cleaning for case-control analysis 14 individuals from CL/CLP group (8 idividuals from cases and 6 individuals from controls) and 8 individuals from CP group (2 individuals from cases and 6 individuals from controls) in Latvian population were removed because of individual missingness treshold of <10%. The overall genotype rate was 100%. After data cleaning for TDT test 2 individuals from CL/CLP group (1 individual from cases and 1 individual from parents as controls) and 13 individuals from CP group (2 individuals from cases and 11 individuals from parents as controls) in Latvian population were removed based on individual missingness treshold of <10%. The overall genotype rate was 100%.

Table 3.13. presents results for case-control comparisons with CL/CLP and CP in Latvian population.

The strongest association with CL/CLP was found for SNP rs2071047, where the allele A was associated with decreased risk (p-value = 0.0087; OR = 0.63; 95% CI = 0.446-0.891) for CL/CLP. Obtained association remained statistically significant after Bonferroni correction. SNP rs17563 showed borderline association (p-value = 0.0178; OR = 0.666; 95% CI = 0.476-0.933) with CL/CLP, which did not remain significant after correction for multiple testing. Allele A was associated with decreased risk for CL/CLP. We did not find any association of analyzed SNPs in *BMP4* gene with isolated CP.

Table 3.13.

BMP4 gene results of case-control analysis associated with non-syndromic CL/CLP and CP in Latvian population

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
CL/CLP									
14	*BMP4*	rs17563	34580722	G/A	0.349	0.446	0.0178	0.666	0.476-0.933
14	*BMP4*	rs2071047	34581611	G/A	0.303	0.408	0.0087	0.63	0.446-0.891
14	*BMP4*	rs1957860	34592555	T/C	0.429	0.446	0.679	0.933	0.672-1.296
CP									
14	*BMP4*	rs17563	34580722	G/A	0.386	0.446	0.354	0.781	0.463-1.318
14	*BMP4*	rs2071047	34581611	G/A	0.386	0.408	0.7322	0.913	0.54-1.541
14	*BMP4*	rs1957860	34592555	T/C	0.529	0.446	0.2019	1.395	0.836-2.328

* Chr - chromosome;
^ SNP - single nucleotide polymorphism;
# Major allele is listed first;
** MAF - minor allele frequency;
^^ OR - odds ratio;
95% CI - 95% confidence interval

Haplotype based association analysis was performed to find any additional possible association in *BMP4* gene with CL/CLP and CP in Latvian population.

In the Table 3.14. results of haplotype analysis associated with CL/CLP in *BMP4* gene are presented.

Table 3.14.

Results of haplotype analysis associated with non-syndeomic CL/CLP in *BMP4* gene

Haplotype	SNP^ 1	SNP 2	SNP 3	Frequency		P value
				Cases	Controls	
*	rs17563	rs2071047	rs1957860	*	*	*
WIN1	A	A	C	0.058	0.115	0.0184
WIN1	G	G	T	0.29	0.236	0.134
WIN1	A	A	T	0.247	0.297	0.1821
WIN1	G	G	C	0.366	0.325	0.2955
WIN1	A	G	T	0.039	0.028	0.4606
*	rs17563	rs2071047	*	*	*	*
WIN1	A	A	*	0.303	0.408	0.0087
WIN1	G	G	*	0.651	0.554	0.0177
WIN1	A	G	*	0.046	0.038	0.6212
*	rs2071047	rs1957860	*	*	*	*
WIN2	A	C	*	0.057	0.114	0.0165
WIN2	G	T	*	0.325	0.261	0.0855
WIN2	A	T	*	0.246	0.294	0.2003
WIN2	G	C	*	0.372	0.332	0.3072

WIN1 - sliding window 1; WIN2 - sliding window 2
^ SNP - single nucleotide polymorphism; * Empty cell

The strongest association with CL/CLP was found for haplotype rs17563-rs2071047 (AA) (p-value = 0.0087) which was associated with decreased risk for the disease. Two additional haplotypes rs2071047-rs1957860 (AC) (p-value = 0.0165) and rs17563-rs2071047-rs1957860) (AAC) (p-value = 0.0184) also were associated with CL/CLP, where both haplotypes showed protective effect for disease.

Haplotype analysis did not show any association with CP and obtained results are shown in Table 3.15.

Table 3.15.

Results of haplotype analysis associated with non-syndromic CP in *BMP4* gene

Haplotype	SNP^ 1	SNP 2	SNP 3	Frequency		p-value
				Cases	Controls	
*	rs17563	rs2071047	rs1957860	*	*	*
WIN1	A	G	T	0	0.027	0.164
WIN1	G	G	C	0.383	0.319	0.2931
WIN1	A	A	T	0.24	0.291	0.3826
WIN1	A	A	C	0.146	0.121	0.5655
WIN1	G	G	T	0.231	0.242	0.8455
*	rs17563	rs2071047		*	*	*
WIN1	A	G	*	0	0.038	0.0972
WIN1	G	G	*	0.614	0.554	0.354
WIN1	A	A	*	0.386	0.408	0.7322
*	rs2071047	rs1957860	*	*	*	*
WIN2	G	C	*	0.381	0.326	0.3674
WIN2	A	T	*	0.238	0.288	0.3991
WIN2	A	C	*	0.147	0.12	0.5233
WIN2	G	T	*	0.233	0.267	0.558

WIN1 - sliding window 1; WIN2 - sliding window 2
^ SNP - single nucleotide polymorphism; * Empty cell

The transmission disequilibrium test was carried out in Latvian non-syndromic cleft lip with or without cleft palate and isolated cleft palate individuals and their parents to identify transmission distortions. We found borderline association between SNP rs1957860 (p value = 0.0455; OR = 3.0; 95% CI = 0.968-9.302) and CP. No association was found for any analyzed markers with CL/CLP. Obtained results are presented in Table 3.16.

Table 3.16.

Transmission distortion results in *BMP4* gene for Latvian CL/CLP and CP individuals

Chr*	Gene	SNP^	Location	Alleles#	Transmitted minor allele count	Untransmitted allele count	p-value	OR^^	95% CI##
CL/CLP									
14	*BMP4*	rs17563	34580722	G/A	16	19	0.6121	0.842	0.433-1.638
14	*BMP4*	rs2071047	34581611	G/A	11	19	0.1441	0.579	0.276-1.217
14	*BMP4*	rs1957860	34592555	T/C	20	21	0.8759	0.952	0.516-1.757
CP									
14	*BMP4*	rs17563	34580722	G/A	4	9	0.1655	0.445	0.127-1.443
14	*BMP4*	rs2071047	34581611	G/A	4	8	0.2482	0.5	0.151-1.66
14	*BMP4*	rs1957860	34592555	T/C	12	4	0.0455	3.0	0.968-9.302

* Chr - chromosome;
^ SNP - single nucleotide polymorphism;
# Major allele is listed first;
^^ OR - odds ratio;
95% CI - 95% confidence interval

3.3.3. *IRF6* gene

We performed case-control analysis in *IRF6* gene seven SNPs for 85 cleft lip and cleft lip with or without cleft palate (CL/CLP) patients, 27 cleft palate (CP) patients and 148 unrelated unaffected individuals as controls in Latvian population. Transmission disequilibrium test was performed for 63 trios (affected sib with both parents), out of all 49 sibs and their parents were divided in CL/CLP group and 14 trios - in CP group. After data quality cleaning for case-control analysis one marker was excluded based on HWE test (p value<0.05), 5 individuals from controls and 1 patient from CP group were removed because of individual missingnes <10%. The overall genotype rate was 100%. After data cleaning for TDT test 1 individual from CP group (1 individual from parents as controls) in Latvian population were removed based on individual missingness treshold of <10%. The overall genotype rate was 100%.

Table 3.17. presents results for case-control comparisons with CL/CLP and CP in Latvian population.

The strongest association with CL/CLP was found for SNP rs658860, where the allele T was associated with decreased risk (p-value = 0.0244×10^{-3}; OR = 0.412; 95% CI = 0.272-0.625) for CL/CLP. Obtained association remained statistically significant after Bonferroni correction. SNP rs642961 showed strong association (p-value = 0.0019; OR = 2.141; 95% CI = 1.315-3.488) with CL/CLP, which also remain significant after correction for multiple testing. Allele G was associated with decreased risk for CLP/CP. Similar results were obtained for rs658860 with CP (p-value = 0.0378×10^{-5}; OR = 0.412; 95% CI = 0.036-0.289), where allele T was also associated with decreased risk for disease and the association reamained significant after Bonferroni correction.

Table 3.17.

IRF6 gene results of case-control analysis associated with non-syndrmic CL/CLP and CP in Latvian population

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
CL/CLP									
1	*IRF6*	rs4844880	207937539	A/T	0.194	0.143	0.1551	1.439	0.87-2.382
1	*IRF6*	rs2013162	208035307	A/C	0.306	0.378	0.1206	0.726	0.485-1.088
1	*IRF6*	rs861019	208042009	G/A	0.429	0.483	0.2714	0.807	0.551-1.183
1	*IRF6*	rs2073487	208043269	C/T	0.312	0.378	0.1548	0.747	0.499-1.117
1	*IRF6*	rs642961	208055893	A/G	0.247	0.133	0.0019	2.141	1.315-3.488
1	*IRF6*	rs658860	208057172	C/T	0.253	0.451	0.0244×10^{-3}	0.412	0.272-0.625
CP									
1	*IRF6*	rs4844880	207937539	A/T	0.289	0.143	0.0096	1.439	1.221-4.806
1	*IRF6*	rs2013162	208035307	A/C	0.481	0.378	0.1614	0.726	0.842-2.765
1	*IRF6*	rs861019	208042009	G/A	0.385	0.483	0.1931	0.807	0.366-1.227
1	*IRF6*	rs2073487	208043269	C/T	0.481	0.378	0.1614	0.747	0.842-2.765
1	*IRF6*	rs642961	208055893	A/G	0.077	0.133	0.2606	2.141	0.186-1.595
1	*IRF6*	rs658860	208057172	C/T	0.077	0.451	0.0378×10^{-5}	0.412	0.036-0.289

* Chr - chromosome; ^ SNP - single nucleotide polymorphism; # Major allele is listed first
** MAF - minor allele frequency; ^^ OR - odds ratio; ## 95% CI - 95% confidence interval

Haplotype based association analysis was performed to find any additional possible association with CL/CLP and CP in Latvian population.

The strongest associations with CL/CLP were found for nine *IRF6* haplotypes - rs642961-rs658860 (GC) (p value= $0.0093x10^{-13}$), rs2073487-rs642961-rs658860 (TGC) (p value= $0.0078x10^{-6}$), rs861019-rs2073487-rs642961-rs658860 (GTGC) (p-value = $0.0131x10^{-6}$), rs2013162-rs861019-rs2073487-rs642961-rs658860 (CGGC) (p-value = $0.0155x10^{-6}$), rs2073487-rs642961-rs658860 (CGC) (p-value = $0.0085x10^{-4}$), rs861019-rs2073487-rs642961-rs658860 (2212) (p-value = $0.0113x10^{-4}$) and rs2013162-rs861019-rs2073487-rs642961-rs658860 (AACGC) (p-value = $0.0114x10^{-4}$), which were associated with an increased risk for CL/CLP, but rs642961-rs658860 (GT) (p-value = $0.0244x10^{-3}$) and rs2013162-rs861019 (CA) (p-value = 0.009), which were associated with lower risk of CL/CLP.

Table 3.18. shows best results of haplotype analysis (p-value ≤0.001) in *IRF6* gene between CL/CLP patients and controls. All results of performed analysis are shown in Appendix 11.

We found very strong association for two *IRF6* haplotypes, where one haplotype rs642961-rs658860 (GT) (p-value = $0.0378x10^{-5}$) was associated with higher risk of this CP and other haplotype rs642961-rs658860 (GC) (p-value = $0.0195x10^{-4}$) was associated with lower risk for CP.

In the Table 3.19. best results of haplotype analysis (p-value ≤0.001) in *IRF6* gene between CP patients and controls are presented. All results of haplotype analysis are shown in Appendix 12.

Table 3.18.

Best results of haplotype analysis in *IRF6* gene associated with non-syndrmic CL/CLP in Latvian population

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs2013162	rs861019	rs2073487	rs642961	rs658860	*	*	*
WIN2	C	G	T	G	C	0.006	0.184	$0.0155x10^{-6}$
WIN2	A	A	C	G	C	0	0.13	$0.0114x10^{-4}$
WIN2	C	A	*	*	*	0.265	0.14	0.0009
*	rs861019	rs2073487	rs642961	rs658860	*	*	*	*
WIN3	G	T	G	C	*	0.006	0.185	$0.0131x10^{-6}$
WIN3	A	C	G	C	*	0	0.13	$0.0113x10^{-4}$
*	rs2073487	rs642961	rs658860	*	*	*	*	*
WIN4	T	G	C	*	*	0.006	0.187	$0.0078x10^{-6}$
WIN4	C	G	C	*	*	0	0.131	$0.0085x10^{-4}$
*	rs642961	rs658860	*	*	*	*	*	*
WIN5	G	C	*	*	*	0.006	0.318	$0.0093x10^{-13}$
WIN5	G	T	*	*	*	0.747	0.549	$0.0244x10^{-3}$

WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4; WIN5 - sliding window 5
^ SNP - single nucleotide polymorphism; * Empty cell

Table 3.19.

The best results of haplotype analysis in *IRF6* gene associated with non-syndromic CP in Latvian population

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs2013162	rs861019	rs2073487	rs642961	rs658860	*	*	*
WIN2	C	G	T	G	C	0	0.20	0.0004
WIN2	C	A	T	G	T	0.058	0.004	0.001
WIN2	A	A	C	G	T	0.481	0.26	0.0013
*	rs861019	rs2073487	rs642961	rs658860	*	*	*	*
WIN3	G	T	G	C	*	0	0.199	0.0004
WIN3	A	C	G	T	*	0.481	0.259	0.0013
*	rs2073487	rs642961	rs658860	*	*	*	*	*
WIN4	T	G	C	*	*	0	0.196	0.0005
WIN4	C	G	T	*	*	0.481	0.255	0.001
*	rs642961	rs658860	*	*	*	*	*	*
WIN5	G	T	*	*	*	0.923	0.549	0.0378×10^{-5}
WIN5	G	C	*	*	*	0	0.318	0.0195×10^{-4}

WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4; WIN5 - sliding window
^ SNP - single nucleotide polymorphism; * Empty cell

We performed transmission disequilibrium test to identify transmission distortions between Latvian non-syndromic CL/CLP/CP individuals and their parents. We found strong association with CP for rs642961 (p-value = 0.0039; OR = 0.091; 95% CI = 0.012-0.704), which remain significant after correction for multiple testing, but borderline association was found between rs658860 (p-value = 0.0067; OR = 0.1; 95% CI = 0.013-0.781) and CP. Both markers were associated with lower risk for CP. The same markers showed significant association with CL/CLP, but they were associated with increased risk for CL/CLP phenotype (rs642961 (p-value = 0.0035; OR = 3.143; 95% CI = 1.343-7.357), rs658860 (p-value = 0.0054; OR = 3.0; 95% CI = 1.275-7.057)) even after Bonferroni correction. Table 3.20. presents all results of transmission distortion in *IRF6* gene for Latvian CL/CLP and CP individuals.

Table 3.20.

Transmission distortion results in *IRF6* gene for Latvian CL/CLP and CP individuals

Chr*	Gene	SNP^	Location	Alleles#	Transmitted minor allele count	Untransmitted allele count	p-value	OR^^	95% CI##
CL/CLP									
1	*IRF6*	rs4844880	207937539	G/A	14	14	1.0	1.0	0.477-2.098
1	*IRF6*	rs2013162	208035307	C/A	13	28	0.0192	0.464	0.241-0.896
1	*IRF6*	rs861019	208042009	G/A	20	25	0.3763	0.8	0.444-1.44
1	*IRF6*	rs2073487	208043269	G/A	13	28	0.0192	0.464	0.241-0.896
1	*IRF6*	rs642961	208055893	T/C	22	7	0.0035	3.143	1.343-7.357
1	*IRF6*	rs658860	208057172	A/G	21	7	0.0054	3.0	1.275-7.057
1	*IRF6*	rs2235371	208030703	T/C	2	3	0.6547	0.667	0.111-3.99
CP									
1	*IRF6*	rs4844880	207937539	G/A	6	3	0.3173	2.0	0.50-7.997
1	*IRF6*	rs2013162	208035307	C/A	10	4	0.1088	2.5	0.784-7.971
1	*IRF6*	rs861019	208042009	G/A	4	4	1.0	1.0	0.25-3.998
1	*IRF6*	rs2073487	208043269	G/A	10	4	0.1088	2.5	0.784-7.971
1	*IRF6*	rs642961	208055893	T/C	1	11	0.0039	0.091	0.012-0.704
1	*IRF6*	rs658860	208057172	A/G	1	10	0.0067	0.1	0.013-0.781
1	*IRF6*	rs2235371	208030703	T/C	0	1	1	NAN	NAN

* Chr - chromosome; ^ SNP - single nucleotide polymorphism; # Major allele is listed first; ^^ OR - odds ratio; ## 95% CI - 95% confidence interval

4. DISCUSSION

Many genes are involved and regulate the development of the craniofacial region. Different growth factors (e.g., FGFs, TGFs, PDGFs, EGFs, BMPs and respective receptors), signaling molecules (e.g., WNT family, SHH and respective receptors), cell adhesion molecules (PVRL1) and transcription factors (e.g., MSX, DLX, LHX, PRRX and BARX family and respective receptors) are encoded by genes which might be involved in the development of non-syndromic cleft lip with or without cleft palate and isolated cleft palate. The human fibroblast growth factors (FGFs) and their cell surface receptors (FGFRs) are a complex family of signaling molecules, that play important roles in a variety of processes of embryogenesis and tissue homeostasis (Itoh and Ornitz, 2004; Chen and Deng, 2005; Dailey et al., 2005; Eswarakumar et al., 2005; Krivicka-Uzkurele et al., 2008). Riley et al. (2007a) performed genome-wide linkage scan in 220 multiplex extended Filipino kindreds for cleft lip with or without cleft palate and identified a novel region at 8p11-23 that is likely to be involved in non-syndromic cleft lip with or without cleft palate. Genes within this region, including *FGFR1* gene, which is localized at 8p12, are considered as possible candidate genes for non-syndromic CL/CLP or CP. Genetic variations in *FGFR1* gene in interaction with non-syndromic CL/CLP or CP have been analyzed in many populations, but observed results are controversial between different populations and races (Riley, 2007a, 2007b; Menezes et al., 2008; Mostowska, 2010; Butali, 2011; Wang, 2011). In our study, after performing the single marker association analysis, one marker in *FGFR1* gene (rs7829058) showed the strongest evidence of association with both non-syndromic cleft lip and cleft lip with cleft palate (p-value = 0.0024×10^{-5}; OR = 7.991; 95% CI = 3.435-18.59) and isolated cleft palate (p-value = 0.0002×10^{-6}; OR = 13.16; 95% CI = 4.93-35.1), which remained significant after Bonferroni correction. We performed haplotype analysis of *FGFR1* gene and observed the association with non-syndromic isolated cleft palate. In another study to enlarge the study population, we pooled together our samples with samples from Estonia and Lithuania. Results of this study showed that SNP rs7829058 is associated with non-syndromic CL/CLP, but the association did not remain significant after Bonferroni correction (p-value = 0.0137; OR = 1.457; 95% CI = 1.079-1.968) (Nikopensius et al., 2011). Single marker association analysis in *FGF1* gene showed that SNP rs34010 (p-value = 0.0002; OR = 0.485; 95%

CI = 0.331-0.71) is associated with protective effect for non-syndromic CL/CLP, but this association did not reach significant results after Bonferroni correction. When we compared isolated cleft palate samples with control samples, we found that SNP rs34016 in *FGF1* gene is associated with increased risk for CP (p-value = 0.006, OR = 2.934; 95% CI = 1.322-6.512), however this association also did not remain significant after correction for multiple comparisons. Haplotype analysis in *FGF1* gene showed positive association with non-syndromic CL/CLP, but not with CP phenotype. Riley et al. (2007a) performed study, where they found that both linkage and association results were positive (recessive multipont HLOD = 1.07) for markers in *FGFR1* gene. The same author in other study (2007b) sequenced the coding regions and performed association testing on 12 genes (*FGFR1*, *FGFR2*, *FGFR3*, *FGF2*, *FGF3*, *FGF4*, *FGF7*, *FGF8*, *FGF9*, *FGF10*, *FGF18*, and *NUDT6*) in population from Iowa and Philippine and used protein structure analyses to predict the function of amino acid variants. They identified few likely disease-causing mutations, including one nonsense mutation (R609X) in *FGFR1* and other missense variants in *FGFR1*, *FGFR2* and *FGFR3* genes. Structural analysis of *FGFR1* variants suggested that identified mutations would impair the function of the proteins through different mechanisms. They also performed SNPs genotyping and found an association between non-syndromic CLP and SNP rs13317 in *FGFR1* gene (p-value = 0.03). The case-control study results in Brazilian population performed by Menezes et al. (2008) partially corroborate the association data, presented by Riley et al. (2007b), in which several genes including *FGFR1* gene, demonstrated a trend for association with non-syndromic CLP with or without dental anomalies. Differencies in the frequencies of the alleles of each polymorphism between cases and controls by each cleft subphenotype were assessed by using OR and 95% CI and they found modest association between SNP rs13317 and right unilateral CLP with tooth agenesis. Mostowska et al. at 2010 published the study, where authors analyzed genes encoding transcription factors such as *FGF10* and *FGFR1* in Polish population. They analyzed two SNPs (rs6987534 and rs328300) in *FGFR1* gene for allelic association, but none of both SNPs were associated with non-syndromic CL/CLP or CP. These markers were analyzed in our study, but we did not find any association between both markers and non-syndromic CL/CLP or CP in our population. Only evidence that these markers could be involved in the etiology for non-syndromic CL/CLP or CP in Latvian population, were results of haplotype analysis, where haplotypes of *FGFR1* gene including both SNPs, showed increased risk for non-syndromic CL/CLP and CP

phenotype. Very similar results were found by Wang et al. at 2011 after testing SNPs in 10 genes coding for fibroblast growth factors and their receptors (including *FGF2* and *FGFR1* gene) in Asian and Maryland case-parent trios ascertained through a child with non-syndromic CL/CLP. They found that *FGFR1* yielded evidence of linkage and association in the TDT, confirming previous evidence. Haplotypes consisting of three SNPs (rs6987534, rs6474354 and rs10958700) in *FGFR1* gene were nominally significant among Asian trios similarly to haplotype based association results found in our study. Negative association results were found by Butali et al. at 2011 after performing genotype association studies and direct sequencing on the *FGFR1* and *FGFR2* genes in Nigerian population. These results can be explained by cleft etiology between different races. The prevalence of CL/CLP/CP in Africa has been reported as relatively lower compared to other populations.

SKI is a proto-oncogene that is required for development of the central nervous system and skeletal muscle, and is involved in specifying selected cranial neural-crest-derived craniofacial structures (Berk et al., 1997). Relatively few of the studies exist on *SKI* gene and its possible role in the development of CL/CLP or CP. Vieira et al. (2005) reported direct sequencing approach to study 20 candidate genes in Philippines for non-syndromic cleft lip with or without cleft palate and the sequencing results suggested that rare point mutations in *FOXE1*, *GLI2*, *JAG2*, *LHX8*, *MSX1*, *MSX2*, *SATB2*, *SPRY2*, *TBX10* and *SKI* gene may be causes of non-syndromic CL/CLP and the linkage disequilibrium data supported a larger, not yet specified, role for variants in or near *MSX2*, *SKI* or *JAG2* genes. To identify genetic variants within the *SKI* gene and investigate the potential association between *SKI* polymorphisms and risk for orofacial clefts, Lu with colleagues (2005) re-sequenced the gene. They identified one novel polymorphism (257C>G) in exon 1, which was associated with the decreased risk (OR = 0.6; 95% CI = 0.3-1.0) for non-syndromic CL/CLP in Californian population. This SNP is located very close to the promoter region so it is possible that this may be in linkage disequilibrium with sequence variants in upstream regulatory regions. In our study we analyzed twenty SNPs in *SKI* gene, which is located at 1q22-q24, for allelic association with non-syndromic CL/CLP or CP. SNP rs16824948 was significantly associated with non-syndromic CL/CLP (p-value = 0.0013×10^{-14}; OR = 6.37; 95% CI = 4.039-1.07) and CP (p-value = 0.0011×10^{-7}; OR = 6.777; 95% CI = 3.577-12.84), where the allele T was associated with increased risk for CL/CLP and CP and this association remained significant after correction for multiple testing. We performed haplotype

analysis and the results of this analysis showed similar results. Unfortunately after pooling data together with Estonians and Lithuanians in our previously published study, we did not observe any significant (p-value ≤ 0.05) association with analyzed markers in *SKI* gene and CL/CLP, but SNP rs12562937 showed borderline association with CP (p-value = 0.0143; OR = 0.534; 95% CI = 0.321-0.889), which did not remain significant after Bonferroni correction (Nikopensius et al., 2010; Nikopensius et al., 2011). Our results support previous positive findings for *SKI* gene role in the etiology of non-syndromic CL/CLP/CP, but additional studies are necessary to replicate obtained results in other populations.

It has been discovered that Wnt signalling pathway plays a crucial role in craniofacial development, and three previously reported studies have concluded that genetic variations in *WNT3* and *WNT9B* genes might be associated with CL/CLP/CP in humans in different populations (Chiquet, 2008; Menezes, 2010; Mostowska, 2012). Twenty-nine single nucleotide polymorphisms in *WNT3* and *WNT9B* genes, located in 17q21, were analyzed in our study for association with non-syndromic CL/CLP and CP in case-control population. One marker in *WNT3* gene (rs11655598) showed the strongest evidence of association with both non-syndromic CL/CLP (p-value = 0.0053×10^{-11}; OR = 5.925; 95% CI = 3.593-9.772) and isolated cleft palate (p-value = 0.0039×10^{-11}; OR = 9.495; 95% CI = 4.879-18.34). This association remained significant after Bonferroni correction. Haplotype based association analysis supports this finding as well. Chiquet et al. (2008) analyzed thirty-eight SNPs in seven *WNT* family genes (*WNT3*, *WNT3A*, *WNT5A*, *WNT7A*, *WNT8A*, *WNT9B* and *WNT11*) in Hispanic and European American population and SNPs in three genes (*WNT3A*, *WNT5A* and *WNT11*) were significantly associated with non-syndromic CL/CLP after correction for multiple testing. Multiple haplotypes in *WNT* family genes were associated with non-syndromic CL/CLP too. Menezes et al. (2010) performed analysis for thirteen SNPs spanning six *WNT* genes (*WNT3*, *WNT3A*, *WNT5A*, *WNT8A*, *WNT9B* and *WNT11*) based on recent publications regarding confirmed associations with non-syndromic cleft lip with or without cleft palate in humans (Chiquet et al., 2008) or in animal models (Juriloff et al., 2005; Juriloff et al., 2006; Lan et al., 2006) to test for association with CL/CLP and CP subphenotypes in Brazilian population. They found that individuals carrying variant alleles in *WNT3* presented an increased risk for "all clefts" (CL/CLP/CP) and cleft lip and cleft lip with or without cleft palate (CL/CLP). SNP rs142167, located in the 5'UTR of *WNT3* gene, showed association with the

phenotype "all clefts" (p-value = 0.0003; OR = 1.61; 95% = 1.29-2.02), cleft lip with palate (CLP) (p-value = 0.001; OR = 1.6; 95% CI = 1.26-2.02) and "unilateral CLP" (p-value = 0.002; OR = 1.65; 95% = 1.27-2.13). Under a nominal value of 0.05, SNP rs9890413 in the same gene also showed an association with "all clefts" (p-value = 0.03; OR = 1.46; 95% CI = 1.12-1.74), with CLP (p-value = 0.02; OR = 1.46; 95% CI = 1.16-1.84) and "unilateral CLP" (p-value = 0.04; OR = 1.46; 95% CI = 1.13-1.89) but SNP rs142167 in *WNT3* was associated with "unsuccessful bilateral" cleft subphenotype (p-value = 0.03; OR = 1.57; 95% CI = 1.17-2.11). The results of the haplotype analysis also supported the associations found for the single SNPs. We analyzed two SNPs (rs111769 and rs2165846) described in previously mentioned studies (Chiquet et al., 2008; Menezes et al., 2010) and SNP rs111769 was associated with non-syndromic CP (p-value = 0.0195; OR = 1.931; 95% CI = 1.105-3.374). This association however did not remain significant after Bonferroni correction. Mostowska et al. (2012) analyzed fourteen SNPs in six *WNT* genes (*WNT3*, *WNT3A*, *WNT5A*, *WNT8A*, *WNT9B* and *WNT11*) and authors found that one *WNT3* gene variant rs3809857 revealed a significant association with the risk of non-syndromic cleft lip and cleft lip with or without cleft palate (CL/CLP) whereas allele T was associated with decreased risk for clefts in Polish population (p-value = 0.015; OR = 0.492; 95% CI = 0.276-0.879). Moreover, haplotype analysis revealed that *WNT3* is significantly associated with non-syndromic CL/CLP. Three SNPs (rs12452064, rs2165846 and rs4968282), analyzed in Polish population, were also included in our study also, but none of them showed any significant association with non-syndromic CP in both Latvian and Polish populations (p-value $\geq$0.05). SNP rs4968282 showed borderline association with non-syndromic CL/CLP (p-value = 0.0444; OR = 0.654; 95% CI = 0.431-0.991) in our population, but observed association did not remain significant after multiple testing. This finding was not confirmed in Polish population. In the Baltic study SNP rs11653738 in *WNT3* gene showed association with CP, but it lost its significancy after correction for multiple testing (p-value = 0.0064; OR = 1.518; 95% CI = 1.123-2.053 (Nikopensius et al., 2010). Two SNPs (rs4968282 and rs1105127) in *WNT9B* gene showed association with non-syndromic CL/CLP (p-value = 0.0013; OR = 0.688; 95% CI = 0.548-0.865 and p-value = 0.0377; OR = 1.239; 95% CI = 1.012-1.518, respectively), which did not remain significant after Bonferroni correction (Nikopensius et al., 2011). Our results further support previous findings that *WNT3* gene is one of the susceptibility genes for non-syndromic CL/CLP/CP in Caucasians.

Linkage and association studies in different populations showed significant association with 19q13 locus also called OFC3 (orofacial 3) locus containing number of following genes, *PVR, PVRL2, BCL3,* and *CLPTM1*, but results were controversial (Stanier and Moore, 2004; Wyszynski et al., 1997; Martinelli et al., 1998; Beaty et al., 2001; Fujita et al., 2004; Morkuniene et al., 2007; Park et al., 2009). In this study, we tested SNPs in *BCL3, CLPTM1, PVR* and *PVRL2* genes in families and individuals from Latvia for association with CL/CLP and CP. *BCL3* polymorhisms were also tested in study involving non-syndromic CL/CLP and CP patients and controls from Brazil. We did not find any significant association after correction for multiple testing of analyzed markers in 19q13 locus and CL/CLP or CP phenotypes compared cases and controls in Latvian population, or between SNPs in *BCL3* gene and non-syndromic CL/CLP or CP in Brazilian population. Only indication of possible association was found between *BCL3* SNP rs8103315 (p-value = 0.0396; OR = 0.245; 95% CI = 0.058-1.04) and CP, and between *BCL3* SNP rs4803750 (p-value = 0.0449; OR = 0.496; 95% CI = 0.247-0.996) and CL/CLP (see Appendix 4), but obtained association did not remain significant after Bonferroni correction. Haplotype analysis in *BCL3* gene showed strong association with CL/CLP in Latvian population and borderline association with CP in Brazilian population. Despite the possitive association between haplotypes and non-syndromic CL/CLP and CP, but no significant association between single marker and CL/CLP or CP, means that haplotypes in *BCL3* gene probably do have some functional effect, which have to be clarified. Such marginal results in Brazilian individuals were not unexpected. A previous study with Brazilian families did not observe any suggestion of transmission disequilibrium between *BCL3* and non-syndromic cleft lip with or without cleft palate (Gaspar et al., 2002). Family based association studies are less sensitive than population based association studies and if any possitive association is found from family studies, it will provide strong evidence. We have also performed transmission desequilibrium test (TDT) in Latvian non-syndromic CL/CLP and CP individuals and their parents in order to identify transmission distortions. Only SNP rs10421283 in *BCL3* gene showed borderline association (p-value = 0.0477) with CL/CLP, but not with CP. These findings corroborate previous studies, where an excess of parental transmission of *BCL3* alleles to cleft probands were detected (Maestri et al., 1997; Park et al., 2009). Warrington et al. (2006) studied 19q13 locus and they found an association between non-syndromic cleft lip with or without cleft palate and the *PVR* gene in two independent populations (Iowa and South America) that remained

significant after correction for multiple testing. We however did not find any association between markers in *PVR* and *PVRL2* genes and non-syndromic CL/CLP or CP in Latvian population, similar to Danish and Italian populations (Warrington et al., 2006; Pezzetti et al., 2007). In the Baltic study two markers in *PVRL2* gene (rs519113 and rs2075642) showed association with CL/CLP, but this association did not remained significant after Bonferroni correction (p-value = 0.0039; OR = 0.702; 95% CI = 0.552-0.894 and p-value = 0.0206; OR = 1.347; 95% CI = 1.046-1.733, respectively) (Nikopensius et al., 2011). SNP rs6859 in *PVRL2* gene and two SNPs (rs5127 and rs16979595) in *CLPTM1* gene showed association with cleft palate, but lost its association after correction for multiple testing (p-value = 0.0472; OR = 1.35; 95% CI = 1.003-1.816, p-value = 0.0146; OR = 1.494; 95% CI = 1.081-2.064, p-value = 0.0288; OR = 1.457; 95% CI = 1.038-2.046, respectively) (Nikopensius et al., 2010). If the 19q13 locus has some impact in the development of non-syndromic CL/CLP or CP then only as a low penetrance or as a modifier locus.

There are few studies reported in humans regarding to *BMP4* gene showing positive association with non-syndromic CL/CLP or CP. Results of meta-analysis of 13 genome scans identified six regions on five chromosomes with HLODs ≥3.2 and one of these regions was at 14q21-25 displaying evidence of linkage with non-syndromic cleft lip with or without cleft palate (Marazita et al., 2004). Based on this discovery, Lin et al. (2008) performed case-control study of *BMP4* gene polymorphisms and found association between 538T/C polymorphism (rs17563) and non-syndromic CL/CLP in Chinese population. The results showed that the 538C allele carriers were associated with a significantly increased risk of non-syndromic CL/CLP compared with the noncarriers (p-value = 0.005; OR = 1.52; 95% CI = 1.13-2.03). There is a study, where mutation analysis of *BMP4* gene have been performed, and it showed significant overrepresentation of *BMP4* mutations in cases with a range of lip and orbicularis oris muscle (OOM) defects and an absence of mutations in more than 500 control samples. These findings support a role for *BMP4* in the pathogenesis of non-syndromic cleft lip with or without cleft palate (Suzuki et al., 2009). Suazo et al. (2010) analyzed the association between *BMP4* gene three SNPs (rs762642, rs2855532 and rs1957860) and non-syndromic CL/CLP in 150 unrealated trios from Chilean population. Obtained results showed that there are no significant transmission distortions for individual SNPs as it was observed for haplotypes rs1957860-rs762642 (T-T (p-value = 0.018) and C-T (p-value = 0.015)). Thus, despite the positive association detected between these

haplotypes and non-syndromic clefts, associated haplotypes probably do not have a functional effect on *BMP4* expression or protein activity, but possibly reflect non-syndromic cleft lip with or without cleft palate susceptibility changes, which are in linkage disequilibrium with these polymorphisms. These findings support a role for *BMP4* in non-syndromic cleft lip with or without cleft palate in the admixed Chilean population. In the present study after performing case-control comparisons, we found association between genetic variations in *BMP4* gene and CL/CLP, but not with CP. The strongest association with CL/CLP was found for SNP rs2071047, which is located in intron 4, where the allele A was associated with decreased risk (p-value = 0.0087; OR = 0.63; 95% CI = 0.446-0.891) for CL/CLP. Obtained association remained statistically significant after Bonferroni correction. SNP rs17563, which is located in exon 5, showed only the borderline association (p-value = 0.0178; OR = 0.666; 95% CI = 0.476-0.933) with CL/CLP. Allele A was associated with decreased risk for CL/CLP. Haplotype analysis showed similar results to association analysis - no association was found between haplotypes in *BMP4* gene and isolated cleft palate phenotype, but two haplotypes showed protective effect for non-syndromic CL/CLP. Transmission disequilibrium test performed to detect any transmission distortions in Latvian trios showed controversial results compared to the single marker association. No association was found between SNPs in *BMP4* gene and CL/CLP as it was described in case-control study, but SNP rs1957860, which is located ~ 6kb downstream of gene, showed borderline association with isolated CP (p-value = 0.0455; OR = 3.0; 95% CI = 0.968-9.302). Our results support previous findings that *BMP4* gene plays significant role in the development of non-syndromic CL/CLP and CP. Obtained results showed that *BMP4* gene could be involved in the development of non-syndromic cleft palate (CP) as a contributor, but it could have protective effect in the susceptibility for non-syndromic cleft lip and cleft lip with or without cleft palate (CL/CLP).

There are many studies regarding to *IRF6* as one of the main genes in the development for non-syndromic CL/CLP/CP. In the present study we found strong association between *IRF6* gene SNPs and non-syndromic CL/CLP and CP. The strongest association with CL/CLP and CP was found for SNPs rs658860 and rs642961, both located ~10-11 kb downstream of the *IRF6* gene. Observed association was strongly confirmed by case-control analysis, haplotype analysis and transmission disequilibrium test. Recent study in Chinese population showed association between *IRF6* gene SNP rs2235371, located in exon 6, where TDT and HHRR (haplotype-based

haplotype relative risk) analysis showed association with non-syndromic cleft lip (CL) (Li et al., 2012). The same SNP was analyzed in our study, but we did not find any association with CL/CLP/CP. These results could be explained by the fact that we did not performed case-control comparisons between cleft lip (CL) individuals and controls. It is possible that SNP rs2235371 could be associated with CL in Latvian population, but sample size in the present study is too small to test it. Classically, cleft lip only and cleft lip with cleft palate are categorized together because these two phenotypes are thought to have the same genetic etiology, whereas isolated cleft palate have different genetic background (Harville, 2005), but obtained results in recent studies suggest that cleft lip only and cleft lip with cleft palate might be separate entities with different etiology and pathogenesis (Jugessur et al., 2011). Similar study was performed in Honduran population (Larrabee et al., 2011), where SNPs rs642961 and rs2235371 were analyzed. They found the association between rs2235371 and non-syndromic CL/CLP in both case-control (p-value = 0.01) and family-based association (p-value = 0.01) studies, but no association was found for rs642961, which is proposed to have potential biological significance to *IRF6* expression and function (Pan et al., 2011). Results obtained in this study, are contrary to ours and another study performed by Shi et al. (2011), where both SNPs were analyzed in Chinese population. Obtained results could be explained by different populations analyzed. Studies, reported previous, suggesting that different populations may be affected by different polymorphisms in *IRF6* gene. In the Baltic study marker rs17389541 showed association with CP (p-value = 0.0006; OR = 1.726; 95% CI = 1.263-2.358), supported by analysis of haplotypes including this polymorphism, but this association did not remain significant after correction for multiple testing. This is a novel implication of *IRF6* in non-syndromic CP susceptibility, but there is a necessarity to replicate obtained results in other populations (Nikopensius et al., 2010).

Failure to replicate an association of SNPs for known cleft genes, such as *MSX1* genes with non-syndromic CL/CLP or CP can be caused by allelic or locus heterogeneity in the etiology of cleft formation. Number of patients and controls analyzed in this study could be to small and number of selected single nucleotide polymorphisms in each gene can be insufficient to achieve full gene coverage. Last, we have analyzed only individual genes, but not interaction between genes and environmental factors, which can also be a very important factor in the development for CL/CLP/CP. Our plans are to start analysis of interaction between genes and

environmenatl factors in very near future.

In summary, our results continue to support the involvement of *FGFR1*, *WNT3*, *SKI*, *BMP4* and *IRF6* genes in non-syndromic CL/CLP and CP in humans and shows possible association between 19q13 locus and non-syndromic CL/CLP and CP. Despite all findings we need to perform additional studies to identify potentially functional variants in these genes and replication studies in different populations not only in Caucasians for genes, which showed modest evidence for association with non-syndromic CL/CLP or CP.

One may argue that almost all investigated SNPs are localized in introns or intergenic regions and do not alter transcription factor binding sites or have any other potentially damaging effect. We should also consider that these SNPs, possibly neutral, might also be in linkage disequilibrium with an etiologic variant, which could explain the results observed in this study. Additionally, the results of many performed GWAS showed the SNPs associated with some disease are considered to be functional.

In conclusion, the results of our study stated that the non-syndromic CL/CLP/CP is very complex malformation and that there are still many undiscovered genes involved in the etiology of this malformation and only few genes have a major role in the development of non-syndromic cleft lip with or without cleft palate and isolated cleft palate.

5. CONCLUSIONS

1. Six hundred and seventy five genetic markers were selected for further genotyping within present study located in selected forty five candidate genes to search for significant relationships with non-syndromic CL/CLP and CP.
2. Case-control analysis showed that genetic variants in *SKI*, *FGFR1*, *WNT3* and *IRF6* genes contributes susceptibility to both non-syndromic cleft lip and cleft lip with or without cleft palate (CL/CLP) and cleft palate only (CP) in Latvian population. *BMP4* gene could have protective effect in the susceptibility for non-syndromic cleft lip and cleft lip with or without cleft palate (CL/CLP).
3. Haplotype analysis showed significant association between haplotypes in *SKI*, *FGFR1*, *WNT3* and *IRF6* genes and non-syndromic CL/CLP and CP, and between haplotypes in *BMP4* and *BCL3* genes and non-syndromic CL/CLP.
4. Results of family based association analysis showed significant association between *IRF6* gene and non-syndromic CL/CLP and CP.
5. Comparative analysis for *BCL3* gene five markers showed association between *BCL3* gene haplotypes and non-syndromic isolated cleft palate in Brazilian population.

6. PUBLICATIONS

Publications

1. Lāce B, Prane I, Piekuse L, Akota I, Barkāne B, Krūmiņa A. Lūpas un/vai aukslēju šķeltņu molekulāri ģenētiskie pētījumi Latvijā. RSU zinātnisko rakstu krājums. 2010; 384-388.
2. Nikopensius T, Jagomägi T, Krjutškov K, Tammekivi V, Saag M, Prane I, Piekuse L, Akota I, Barkane B, Krumina A, Ambrozaitytė L, Matulevičienė A, Kučinskienė ZA, Lace B, Kučinskas V, Metspalu A. Genetic variants in COL2A1, COL11A2, and IRF6 contribute risk to nonsyndromic cleft palate. Birth Defects Res A Clin Mol Teratol. 2010; 88(9): 748-756.
3. Nikopensius T, Kempa I, Ambrozaitytė L, Jagomägi T, Saag M, Matulevičienė A, Utkus A, Krjutškov K, Tammekivi V, Piekuse L, Akota I, Barkane B, Krumina A, Klovins K, Lace B, Kučinskas V, Metspalu A. Variation in FGF1, FOXE1, and TIMP2 genes is associated with nonsyndromic cleft lip with or without cleft palate. Birth Defects Res A Clin Mol Teratol. 2011; 91(4): 218-225.
4. Lace B, Kempa I, Piekuse L, Grinfelde I, Klovins J, Pliss L, Krumina A, Vieira AR. Association Studies of Candidate Genes and Cleft Lip and Palate Taking into Consideration Geographic Origins. Eur J Oral Sci. 2011: 119(6): 413-417.
5. Lace B, Kempa I, Klovins J, Stavusis J, Krumina A, Akota I, Barkane B, Vieira AR, Nagle E, Grinfelde I, Maulina I. BCL3 gene role in facial morphology. Birth Defects Res A Clin Mol Teratol. 2012; 94(11): 918-924.
6. Letra A, Fakhouri W, Fonseca RF, Menezes R, Kempa I, Prasad JL, McHenry TG, Lidral AC, Moreno L, Murray JC, Daack-Hirsch S, Marazita ML, Castilla EE, Lace B, Orioli IM, Granjeiro JM, Schutte BC, Vieira AR. Interaction between IRF6 and TGFA genes contribute to the risk of nonsyndromic cleft lip/palate. PLoS One. 2012; 7(9): e45441. doi: 10.1371/journal.pone.0045441.

Abstracts

1. Prane I, Lace B, Piekuse L, Barkane B, Akota I, Krumina A. BCL3 gene association analysis with nonsyndromic cleft lip and/or cleft palate in Latvia. *European Journal of Human Genetics*. 2007; 15(1): 240.

2. Prane I, Lace B, Piekuse L, Klovins J, Barkane B, Akota I, Krumina A. Intron region importance of BCL3 gene in the nonsyndromic cleft and/or celft palate. *European Journal of Human Genetics*. 2008; 16(2): 300.
3. Prane I, Lace B, Vieira AR, Akota I, Barkane B, Krumina A. Significant association between IRF6 gene and nonsyndromic cleft lip with or without cleft palate in Latvian population. *European Journal of Human Genetics*. 2009; 17(2): 237-238.
4. Prane I, Lace B, Vieira AR, Letra A, Menezes R, Barkane B, Akota I, Krumina A. *IRF6 gene as a genetic factor in the development of cleft lip with or without cleft palate*. Stomatologija. 2009; 6: 20.
5. Prane I, Piekuse L, Akota I, Barkane B, Krumina A, Lace B. Polymorphisms in WNT family genes are associated with nonsyndromic cleft lip and palate. *European Journal of Human Genetics*. 2010; 18(1): 248.
6. Prane I, Piekuse L, Akota I, Barkāne B, Krūmiņa A, Lāce B. New concepts in genetics of nonsyndromic orofacial clefts. The 7th Congress of the Baltic Association for Maxillofacial and Plastic Surgery, May 20-22, Riga, Latvia, 2010.
7. Lāce B, Micule I, Grīnfelde I, Prane I, Piekuse L, Krūmiņa A, Barkāne B, Akota I. Identified risk factors of cleft lip and/or palate development. The 7th Congress of the Baltic Association for Maxillofacial and Plastic Surgery, May 20 – 22, Riga, Latvia, 2010.
8. Lāce B, Pliss L, Pelnēna I, Prane I. Our common roots: mitochondrial haplotype analysis in patients with nonsyndromic cleft lip and/or palate. The 7th Congress of the Baltic Association for Maxillofacial and Plastic Surgery, May 20 – 22, Riga, Latvia, 2010.
9. Kempa I, Piekuse L, Akota I, Barkane B, Krumina A, Lace B. New concepts in genetics of nonsyndromic orofacial clefts. 4. starptautiskā jauno zinātnieku skola molekulārajā ģenētikā "Genomika un šūnas bioloģija", Moscow, Russia, 2010; 230.
10. Kempa I, Piekuse L, Klovins J, Barkane B, Akota I, Krumina A, Lace B. Polymorphisms in FGF1 gene are associated with nonsyndromic cleft lip and palate. *European Journal of Human Genetics*. 2011; 19(2): 307.
11. Kempa I, Martinkevica O, Klovins J, Akota I, Barkane B, Krumina A, Lace B. BCL3 gene polymorphisms and nonsyndromic cleft lip with or without cleft

palate. *International Journal of Experimental and Clinical Pathophysiology and Drug Research*. 2011; 25(3): 536.

12. Krasone K, Lāce B, Kempa I, Piekuse L, Akota I, Barkāne B, Care R. Hipodontijas ģenētiskie pētījumi indivīdiem ar lūpu un/vai aukslēju šķeltni Latvijā. RSU Zinātniskā konference, Tēzes 2012; 320.
13. Kempa I, Akota I, Barkane B, Krumina A, Klovins J, Lace B. Asssociation of IRF6 gene polymorphisms with nonsyndromic cleft lip with or without cleft palate. *European Journal of Human Genetics*. 2012; 20(1): 234.
14. Kempa I, Piekuse L, Klovins J, Barkane B, Akota I, Krumina A, Lace B. Polymorphisms in FGF1 gene are associated with nonsyndromic cleft lip and palate. The V-th International school on molecular genetics for young scientists "Variability of the genome", Abstract book, 2012; 8.
15. Kempa I, Akota I, Barkane B, Krumina A, Klovins J, Lace B. Association of IRF6 gene polymorphisms with non-syndromic cleft lip with or without cleft palate. The 3rd Open Gene Young investigator workshop in Baltic region, Abstract book, 2012; 31-32.
16. Kempa I, Piekuse L, Akota I, Barkane B, Krumina A, Stavusis J, Klovins J, Lace B. Genetic variants in SKI, FGFR1, WNT3, IRF6 and BMP4 genes are associated with risk for non-syndromic CL/CLP and/or CP in Latvian population. *European Journal of Human Genetics*. 2013 (accepted).

Approbation

1. Pre-defence of the thesis was held in joint meeting of Department of Biology and Microbiology, Rīga Stradiņš University, Institute of Stomatology, Rīga Stradiņš University and Latvian Association of Human Genetics at April 23, 2012, Stomatology Institute, Rīga Stradiņš University, Riga, Latvia.
2. Kempa I, Martinkevica O, Klovins J, Akota I, Barkane B, Krumina A, Lace B. BCL3 gene polymorphisms and nonsyndromic cleft lip with or without cleft palate. 9th European Craniofacial Congress, 14.09.2011.-17.09.2011., Salzburg, Austria.
3. Kempa I, Akota I, Barkāne B, Krūmiņa A, Lāce B. "Ģenētisko faktoru loma nesindromālo lūpas un/vai aukslēju šķeltņu attīstībā Latvijas populācijā". LZP Sadarbības projekta seminārs, 03.03.2011., Rīga, Latvija.

4. Prane I, Piekuse L, Akota I, Barkāne B, Krūmiņa A, Lāce B. „New concepts in genetics of nonsyndromic orofacial clefts". The 7th Congress of the Baltic Association for Maxillofacial and Plastic Surgery, 20.05.2010.-22.05.2010., Rīga, Latvija.

ACKNOWLEDGEMENTS

I would like to thank to everyone who participated in the process of this thesis and made it possible.

I wish to thank to my both scientific supervisors, dr. med. Baiba Lāce and dr. biol. Jānis Kloviņš for giving me the chance to work in their team and grow up with them as a scientist. From my deepest heart I would like to thank to Baiba for guiding me and for the time she spent reading and correcting my thesis. I would like to thank to Jānis for advices in data statistical analysis.

I am very gratefull to my colleagues from Rīga Stradiņš University and Latvian Biomedical Research and Study Centre, especially to my colleague and dear friend Linda Piekuse for beeing next to me in difficult moments.

I wish to thank to our colleagues from Riga Cleft Lip and Palate Centre, Institute of Stomatology, Rīga Stradiņš University for sample collection.

I would like to thank to all patients and their families for participating in this research because without them this thesis would not be possible.

My deepest gratitude goes to my family, especially to my mom who always supports me and accepts my decisions. Finally, very special gratitude goes to my husband for letting me to fulfill my dreams and for allowing to live this crazy lifestyle called science.

REFERRENCES

1. Akota I, Barkane B, Grasmane N. Iedzimto sejas šķeltņu biežuma rādītāji Latvijā no 1960. līdz 1997. gadam. *Proc Latv Med Acad.* 2001; 166-170.
2. Amos C, Gasser D, Hecht JT. Nonsyndromic cleft lip with or without cleft palate. New BCL3 information. *Am J Hum Genet.* 1996a; 59(3): 743-744.
3. Amos C, Stein J, Mulliken JB et al. Nonsyndromic cleft lip with or without cleft palate: erratum. *Am J Hum Genet.* 1996b; 59: 744.
4. Ashique AM, Fu K, Richman JM. Endogenous bone morphogenetic proteins regulate outgrowth and epithelial survival during avian lip fusion. *Development.* 2002; 129(19): 4647-4660.
5. Beaty TH, Wang H, Hetmanski JB et al. A case-control study of nonsyndromic oral clefts in Maryland. *Ann Epidemiol.* 2001; 11(6): 434-442.
6. Berk M, Desai SY, Heyman HC, Colmenares C. Mice lacking the SKI proto-oncogene have defects in neurulation, craniofacial patterning, and skeletal muscle development. *Genes Dev.* 1997; 11(16): 2029-2039.
7. Bille C, Pedersen DA, Andersen AM, Mansilla MA, Murray JC, Christensen K, Ballard JL, Gorman EB, Cabrera RM, Finnell RH. Autoantibodies to folate receptor alpha during early pregnancy and risk of oral clefts in Denmark. *Pediatric Res.* 2010; 67(3): 274-279.
8. Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeida de Assis N, Alblas MA, Barth S, Freudenberg J, Lauster C, Schmidt G, Scheer M, Braumann B, Bergé SJ, Reich RH, Schiefke F, Hemprich A, Pötzsch S, Steegers-Theunissen RP, Pötzsch B, Moebus S, Horsthemke B, Kramer FJ, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M, Nöthen MM, Mangold E. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nat Genet.* 2009; 41(4): 473-477.
9. Blanco R, Suazo J, Santos JL, Paredes M, Sung H, Carreño H, Jara L. Association between 10 microsatellite markers and nonsyndromic cleft lip palate in the Chilean population. *Cleft Palate Craniofac J.* 2004; 41(2): 163-167.
10. Boyles AL, Wilcox AJ, Taylor JA, Shi M, Weinberg CR, Meyer K, Fredriksen A, Ueland PM, Johansen AM, Drevon CA, Jugessur A, Trung TN, Gjessing HK, Vollset SE, Murray JC, Christensen K, Lie RT. Oral facial clefts and gene

polymorphisms in metabolism of folate/one- carbon and vitamin A: a pathway-wide association study. *Genet Epidemiol.* 2009; 33(3): 247-255.

11. Brugmann SA, Goodnough LH, Gregorieff A, Leucht P, ten Berge D, Fuerer C, Clevers H, Nusse R, Helms JA. Wnt signaling mediates regional specification in the vertebrate face. *Development.* 2007; 134(18): 3283-3295.
12. Butali A, Mossey PA, Adeyemo WL, Jezewski PA, Onwuamah CK, Ogunlewe MO, Ugboko VI, Adejuyigbe O, Adigun AI, Abdur-Rahman LO, Onah II, Audu RA, Idigbe EO, Mansilla MA, Dragan EA, Petrin AL, Bullard SA, Uduezue AO, Osaguona AO, Olasoji HO, Ligali TO, Kejeh BM, Iseh KR, Olaitan PB, Adebola AR, Efunkoya E, Adesina OA, Oluwatosin OM, Murray JC. Genetic studies in the nigerian population implicate an MSX1 mutation in complex oral facial clefting disorders. *Cleft Palate Craniofac J.* 2011; 48(6): 646-653.
13. Carreño H, Suazo J, Paredes M, Solá J, Valenzuela J, Blanco R. Association between cleft lip/palate phenotype and non syndrome microsatellite markers located in 6p, 17q and 19q. *Rev Med Chil.* 2002; 130(1): 35-44.
14. Carter CO. Genetics of common disorders. *Br Med Bull.* 1969; 25(1):52-57.
15. Carter CO. Genetics of common single malformations. *Br Med Bull.* 1976; 32(1):21-6.
16. Chen L, Deng CX. Roles of FGF signaling in skeletal development and human genetic diseases. *Front Biosci.* 2005; 10: 1961-1976.
17. Chiquet BT, Blanton SH, Burt A et al. Variation in WNT genes is associated with non-syndromic cleft lip with or without cleft palate. *Hum Mol Genet.* 2008; 17(14): 2212-2218.
18. Christensen K, Fogh-Andersen P. Cleft lip (+/- cleft palate) in Danish twins, 1970-1990. *Am J Med Genet.* 1993 1;47(6): 910-916.
19. Christensen K, Juel K, Herskind AM, Murray JC. Long term follow up study of survival associated with cleft lip and palate at birth. *BMJ.* 2004; 328(7453): 1405.
20. Christensen K, Mitchell LE. Familial recurrence-pattern analysis of nonsyndromic isolated cleft palate - a Danish Registry study. *Am J Hum Genet.* 1996; 58(1): 182-190.
21. Chung CS, Bixler D, Watanabe T, Koguchi H, Fogh- Andersen P. Segregation analysis of cleft lip with or without cleft palate: a comparison of Danish and Japanese data. *Am JHum Genet.* 1986; 39: 603-611.

22. Chung KS, Park HH, Ting K et al. Modulated expression of type X collagen in the Meckel's cartilage with different developmental fates. *Dev Biol.* 1995; 170: 387-396.
23. Dailey L, Ambrosetti D, Mansukhani A, Basilico C. Mechanisms underlying differential responses to FGF signaling. *Cytokine Growth Factor Rev.* 2005; 16(2): 233-247.
24. Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet. 2011; 12(3):167-178.
25. Deroo LA, Wilcox AJ. First-Trimester Maternal Alcohol Consumption and the Risk of Infant Oral Clefts in Norway: A Population-based Case-Control Study. *Am I Epidemiol.* 2008; 168(6): 638-646.
26. Dodé C, Levilliers J, Dupont J-M, De Pape A, Le Dû N et al. FGFR1 loss-of-function mutations cause autosomal dominant Kallmann syndrome. *Nat Genet.* 2003; 33(4): 463-465.
27. Engert JC, S. Servaes, P. Sutrave, S.H. Hughes, N. Rosenthal, Activation of a muscle-specific enhancer by the SKI proto-oncogene. *Nucleic Acids Res.* 1995; 23(15): 2988-2994.
28. Eswarakumar VP, Lax I, Schlessinger J. Cellular signaling by fibroblast growth factor receptors. *Cytokine Growth Factor Rev.* 2005; 16(2): 139-149.
29. Falconer DS. The inheritance of liability to certain diseases estimated from the incidence among relatives. *Ann Hum Genet.* 1965; 29: 51-76.
30. Ferguson MW. Palate development. *Development.* 1988; 103: 41-60.
31. Fitz Patrick D, Farrall M. An estimation of the number of susceptibility loci for isolated cleft palate. *J Craniofac Genet Dev Biol.* 1993; 13: 230-235.
32. Fogh-Andersen P. Inheritance of harelip and cleft palate. Munksgaard, Copenhagen. 1942.
33. Fraser FC. The multifactorial/threshold concept - uses and misuses. *Teratology.* 1976; 14(3): 267-280.
34. Fujita H, Nagata M, Ono K, Okubo H, Takagi R. Linkage analysis between BCL3 and nearby genes on 19q13.2 and non-syndromic cleft lip with or without cleft palate in multigenerational Japanese families. *Oral Dis.* 2004; 10(6): 353-359.

35. Gaspar DA, Matioli SR, Pavanello RC et al. Evidence that BCL3 plays a role in the etiology of nonsyndromic oral clefts in Brazilian families. *Genet Epidemiol.* 2002; 23(4): 364-374.
36. Gavin BJ, McMahon JA, McMahon AP. Expression of multiple novel Wnt-1/int-1-related genes during fetal and adult mouse development. *Genes Dev.* 1990; 4(12B): 2319-2332.
37. Harville EW, Wilcox AJ, Lie RT, Vindenes H, Abyholm F. Cleft lip and palate versus cleft lip only: are they distinct defects? *Am J Epidemiol.* 2005; 162(5): 448-453.
38. Hinrichsen K. The early development of morphology and patterns of the face in the human embryo. *Adv Anat Embryol Cell Biol.* 1985; 98: 1-79.
39. Hecht JT, Yang P, Michels VV, Buetow KH. Complex segregation analysis of nonsyndromic cleft lip and palate. *Am J Hum Genet.* 1991; 49: 674-681.
40. http://www.eurocran.org
41. http://www.ncbi.nlm.nih.gov
42. http://www.who.int
43. Itoh N, Ornitz DM. Evolution of the Fgf and Fgfr gene families. *Trends Genet.* 2004; 20(11): 563-569.
44. Jiang R, Bush JO, Lidral AC. Development of the upper lip: morphogenetic and molecular mechanisms. *Dev Dyn.* 2006; 235(5): 1152-1166.
45. Jentink J, Loane MA, Dolk H, Barisic I, Garne E, Morris JK, de Jong-van den Berg LT. EUROCAT. Antiepileptic Study Working Group. Valproic acid monotherapy in pregnancy and major congenital malformations. *N Engl J Med.* 2010; 362(23): 2185-2193.
46. Jezewski PA, Vieira AR, Nishimura C, Ludwig B, Johnson M, O'Brien SE, Daack-Hirsch S, Schultz RE, Weber A, Nepomucena B, Romitti PA, Christensen K, Orioli IM, Castilla EE, Machida J, Natsume N, Murray JC. Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. *J Med Genet.* 2003; 40(6): 399-407.
47. Johnson C, Little J. Folate intake, markers of folate status and oral clefts: is the evidence converging. *Int J Epidemiol.* 2008; 37(5):1041-1058.
48. Jones MC. Etiology of facial clefts: prospective evaluation of 428 patients. *Cleft Palate J.* 1988; 25(1): 16-20.
49. Jugessur A, Farlie PG, Kilpatrick N. The genetics of isolated orofacial clefts:

from genotypes to subphenotypes. *Oral Dis*. 2009; 15(7): 437-53.

50. Jugessur A, Shi M, Gjessing HK, Lie RT, Wilcox AJ, Weinberg CR, Christensen K, Boyles AL, Daack-Hirsch S, Nguyen TT, Christiansen L, Lidral AC, Murray JC. Fetal genetic risk of isolated cleft lip only versus isolated cleft lip and palate: a subphenotype analysis using two population-based studies of orofacial clefts in Scandinavia. *Birth Defects Res A Clin Mol Teratol*. 2011; 91(2): 85-92.
51. Juriloff DM, Harris MJ, Brown CJ. Unravelling the complex genetics of cleft lip in the mouse model. *Mamm Genome*. 2001; 12(6): 426-35.
52. Juriloff DM, Harris MJ, Dewell SL, Brown CJ, Mager DL, Gagnier L, Mah DG. Investigations of the genomic region that contains the clf1 mutation, a causal gene in multifactorial cleft lip and palate in mice. *Birth Defects Res A Clin Mol Teratol*. 2005; 73(2): 103-113.
53. Juriloff DM, Harris MJ, Dewell SL. A digenic cause of cleft lip in A-strain mice and definition of candidate genes for the two loci. *Birth Defects Res A Clin Mol Teratol*. 2004; 70(8): 509-518.
54. Kayserili H, Wollnik B, Güven G, Emiroğlu MU, Başerer N, Uyguner ZO. A novel homozygous COL11A2 deletion causes a C-terminal protein truncation with incomplete mRNA decay in a Turkish patient. *Am J Med Genet A*. 2011; 155A(1): 180-185.
55. Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, de Lima RL, Daack-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, Lammer EJ, Aylsworth AS, Ardinger HH, Lidral AC, Pober BR, Moreno L, Arcos-Burgos M, Valencia C, Houdayer C, Bahuau M, Moretti-Ferreira D, Richieri-Costa A, Dixon MJ, Murray JC. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nat Genet*. 2002; 32(2): 285-289.
56. Kriens O. LAHSHAL: An easy clinical system of cleft lip alveolus and palate documentation in Kriens O. Proceedings of the advanced workshop "What is a cleft?" Stuttgart G, *Thieme* 1989.
57. Krivicka-Uzkurele B, Pilmane M, Akota I. Barx1. growth factors and apoptosis in facial tissue of children with clefts. *Stomatologija*. 2008; 10(2): 62-66.
58. Krjutskov K, Andreson R, Mägi R, Nikopensius T, Khrunin A, Mihailov E, Tammekivi V, Sork H, Remm M, Metspalu A. Development of a single tube

640-plex genotyping method for detection of nucleic acid variations on microarrays. *Nucleic Acids Res*. 2008; 36(12): e75.

59. Lammer EJ, Shaw GM, Iovannisci DM, Van Waes J, Finnell RH. Maternal smoking and the risk of orofacial clefts: Susceptibility with NAT1 and NAT2 polymorphisms. *Epidemiology*. 2004; 15(2): 150-156.
60. Larrabee YC, Birkeland AC, Kent DT, Flores C, Su GH, Lee JH, Haddad J Jr. Association of common variants, not rare mutations, in IRF6 with nonsyndromic clefts in a Honduran population. *Laryngoscope*. 2011; 121(8): 1756-1759.
61. Li M, Zhu W, Wang Y, Guo J, Li S, Li Y. [Association between polymorphism of IRF6 rs2235371 locus and nonsyndromic cleft lip with or without cleft palate]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi*. 2012; 29(2): 149-154.
62. Lin JY, Chen YJ, Huang YL, Tang GP, Zhang L, Deng B, Li M, Ma H, Luan RS. Association of bone morphogenetic protein 4 gene polymorphisms with nonsyndromic cleft lip with or without cleft palate in Chinese children. *DNA Cell Biol*. 2008; 27(11): 601-605.
63. Little J, Cardy A, Munger RG. Tobacco smoking and oral clefts: a meta-analysis. *Bull World Health Organ*. 2004; 82(3):213-218.
64. Liu W, Sun X, Braut A, Mishina Y, Behringer RR, Mina M, Martin JF: Distinct functions for Bmp signaling in lip and palate fusion in mice. *Development*. 2005; 132(6): 1453-1461.
65. López-Camelo JS, Castilla EE, Orioli IM. INAGEMP (Instituto Nacional de Genética Médica Populacional); ECLAMC (Estudio Colaborativo Latino Americano de Malformaciones Congénitas). Folic acid flour fortification: impact on the frequencies of 52 congenital anomaly types in three South American countries. *Am J Hum Genet A*. 2010; 152A(10): 2444-2458.
66. Lu W, Volcik K, Zhu H, Wen S, Shaw GM, Lammer EJ, Finnell RH. Genetic variation in the proto-oncogene SKI and risk for orofacial clefting. *Mol Genet Metab*. 2005; 86(3): 412-416.
67. Luo K. Negative regulation of BMP signaling by the ski oncoprotein. *J Bone Joint Surg Am*. 2003; 85-A(3): 39-43.
68. Maestri NE, Beaty TH, Hetmanski J et al. Application of transmission disequilibrium tests to nonsyndromic oral clefts: including candidate genes and environmental exposures in the models. *American Journal of Medical Genetics*. 1997; 73(3): 337-344.

69. Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, Reutter H, de Assis NA, Chawa TA, Mattheisen M, Steffens M, Barth S, Kluck N, Paul A, Becker J, Lauster C, Schmidt G, Braumann B, Scheer M, Reich RH, Hemprich A, Pötzsch S, Blaumeiser B, Moebus S, Krawczak M, Schreiber S, Meitinger T, Wichmann HE, Steegers-Theunissen RP, Kramer FJ, Cichon S, Propping P, Wienker TF, Knapp M, Rubini M, Mossey PA, Hoffmann P, Nöthen MM. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. *Nat Genet*. 2010; 42(1): 24-26.

70. Mangold E, Reutter H, Birnbaum S, Walier M, Mattheisen M, Henschke H, Lauster C, Schmidt G, Schiefke F, Reich RH, Scheer M, Hemprich A, Martini M, Braumann B, Krimmel M, Opitz C, Lenz JH, Kramer FJ, Wienker TF, Nöthen MM, Diaz Lacava A. Genome-wide linkage scan of nonsyndromic orofacial clefting in 91 families of central European origin. *Am J Med Genet A*. 2009; 149A(12): 2680-2694.

71. Marazita ML, Field LL, Cooper ME, Tobias R, Maher BS, Peanchitlertkajorn S, Liu YE. Nonsyndromic cleft lip with or without cleft palate in China: assessment of candidate regions. *Cleft Palate Craniofac J*. 2002; 39: 149-156.

72. Marazita ML, Goldstein AM, Smalley SL, Spence MA. Cleft lip with or without cleft palate: reanalysis of a three-generation family study from England. *Genet Epidemiol*. 1986; 3(5): 335-342.

73. Marazita ML, Lidral AC, Murray JC et al. Genome scan, fine-mapping, and candidate gene analysis of non-syndromic cleft lip with or without cleft palate reveals phenotype-specific differences in linkage and association results. *Hum Hered*. 2009; 68(3): 151-170.

74. Marazita ML, Murray JC, Lidral AC, Arcos-Burgos M, Cooper ME, Goldstein T, Maher BS, Daack-Hirsch S, Schultz R, Mansilla MA, Field LL, Liu YE, Prescott N, Malcolm S, Winter R, Ray A, Moreno L, Valencia C, Neiswanger K, Wyszynski DF, Bailey-Wilson JE, Albacha-Hejazi H, Beaty TH, McIntosh I, Hetmanski JB, Tunçbilek G, Edwards M, Harkin L, Scott R, Roddick LG. Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35. *Am J Hum Genet*. 2004; 75(2): 161-173.

75. Marazita ML, SpenceMA, MelnickM. Genetic analysis of cleft lip with or without cleft palate in Danish kindreds. *Am JHum Genet*. 1984; 19: 9-18.

76. Marie PJ, Coffin JD, Hurley MM. FGF and FGFR signalingin chondro-dysplasias and craniosynostosis. *J Cell Biochem*. 2005; 96: 888-896.

77. Martinelli M, Scapoli L, Pezzetti F et al. Suggestive linkage between markers on chromosome 19q13.2 and nonsyndromic orofacial cleft malformation. *Genomics*. 1998; 51(2): 177-181.

78. McKeithan TW, Rowley JD, Shows TB, Diaz MO. Cloning of the chromosome translocation breakpoint junction of the t(14;19) in chronic lymphocytic leukemia. *Proc Natl Acad Sci.* 1987; 84(24): 9257-9260.

79. Melkoniemi M, Koillinen H, Männikkö M, Warman ML, Pihlajamaa T, Kääriäinen H, Rautio J, Hukki J, Stofko JA, Cisneros GJ, Krakow D, Cohn DH, Kere J, Ala-Kokko L. Collagen XI sequence variations in nonsyndromic cleft palate, Robin sequence and micrognathia. *Eur J Hum Genet*. 2003; 11(3): 265-270.

80. Melnick M, Bixler D, Fogh-Andersen P, Conneally PM. Cleft lip+/-cleft palate: an overview of the literature and an analysis of Danish cases born between 1941 and 1968. *Am J Med Genet*. 1980; 6(1): 83-97.

81. Menezes R, Letra A, Ruff J, Granjeiro JM, Vieira AR. Studies of genes in the FGF signaling pathway and oral clefts with or without dental anomalies. *Am J Med Genet A*. 2008; 15; 146A(12): 1614-1617.

82. Menezes R, Letra A, Kim AH, Küchler EC, Day A, Tannure PN, Gomes da Motta L, Paiva KB, Granjeiro JM, Vieira AR. Studies with Wnt genes and nonsyndromic cleft lip and palate. *Birth Defects Res A Clin Mol Teratol*. 2010; 88(11): 995-1000.

83. Mitchell LE, Risch N. Mode of inheritance of nonsyndromic cleft lip with or without cleft palate: a reanalysis. *Am J Hum Genet*. 1992; 51(2): 323-332.

84. Mooney MP, Siegel MI, editors: Understanding craniofacial anomalies: The etiopathogenesis of craniosynostoses and facial clefting. New York, *Wiley-Liss, Inc*; 2002.

85. Morkuniene A, Steponaviciute D, Utkus A, Kucinskas V. Few associations of candidate genes with nonsyndromic orofacial clefts in the population of Lithuania. *J Appl Genet.* 2007; 48(1): 89-91.

86. Mossey P, Little J, Munger RG, Dixon MJ, Shaw WC. Cleft lip and palate. *Lancet*. 2009; 374(9703): 1773-1785.

87. Mossey P, Little J. Addressing the challenges of cleft lip and palate research in India. *Indian J Plast Surg*. 2009; 42: 9-18.

88. Mostowska A, Hozyasz KK, Wojcicki P, Biedziak B, Paradowska P, Jagodzinski PP. Association between genetic variants of reported candidate genes or regions and risk of cleft lip with or without cleft palate in the polish population. *Birth Defects Res A Clin Mol Teratol*. 2010; 88(7): 538-545.

89. Mostowska A, Hozyasz KK, Biedziak B, Wojcicki P, Lianeri M, Jagodzinski PP. Genotype and haplotype analysis of WNT genes in non-syndromic cleft lip with or without cleft palate. *Eur J Oral Sci*. 2012; 120(1): 1-8.

90. Munger RG, Tamura T, Johnston KE, Feldkamp ML, Pfister R, Carey JC. Plasma zinc concentrations of mothers and the risk of oral clefts in their children in Utah. *Birth Defects Res A Clin Mol Teratol*. 2009; 85(2): 151-155.

91. Murray JC. Face facts: genes, environment, and clefts. *Am J Hum Genet*. 1995; 57(2): 227-232.

92. Murray JC: Gene/environment causes of cleft lip and/or palate. *Clin Genet*. 2002; 61(4): 248-256.

93. Nagase T, Mizuguchi G, Nomura N, Ishizaki R, Ueno Y, Ishii S. Requirement of protein co-factor for the DNA-binding function of the human SKI proto-oncogene product, *Nucleic Acids Res*. 1990; 18 337-343.

94. Nanci A. Ten Cate's Oral Histology - Development, Structure and *Function,* 6th edn. *Los Angelos: Mosby,* 2003.

95. Niemann S, Zhao C, Pascu F et al. Homozygous WNT3 mutation causes tetra-amelia in a large consanguineous family. *Am J Hum Genet.* 2004; 74(3): 558-563.

96. Pan Y, Ma J, Zhang W, Du Y, Niu Y, Wang M, Zhang Z, Wang L. IRF6 polymorphisms are associated with nonsyndromic orofacial clefts in a Chinese Han population. *Am J Med Genet A*. 2010; 152A(10): 2505-211.

97. Park BY, Sull JW, Park JY, Jee SH, Beaty TH. Differential parental transmission of markers in BCL3 among Korean cleft case-parent trios. *J Prev Med Public Health*. 2009; 42(1): 1-4.

98. Pezzetti F, Palmieri A, Martinelli M, Scapoli L, Arlotti M, Baciliero U, Padula E, Carinci P, Caramelli E, Carinci F. Linkage disequilibrium analysis of two genes mapping on OFC3: PVR and PVRL2. *Eur J Hum Genet*. 2007; 15(9): 992-994.

99. Porter F. Cholesterol precursors and facial clefting. *J Clin Invest*. 2006; 116(9): 2322-2325.

100. Pullat J, Metspalu A. Arrayed primer extension reaction for genotyping on oligonucleotide microarray. *Methods Mol Biol*. 2008; 444: 161-167.

101. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3): 559-575.

102. Rahimov F, Marazita ML, Visel A, Cooper ME, Hitchler MJ, Rubini M, Domann FE, Govil M, Christensen K, Bille C, Melbye M, Jugessur A, Lie RT, Wilcox AJ, Fitzpatrick DR, Green ED, Mossey PA, Little J, Steegers-Theunissen RP, Pennacchio LA, Schutte BC, Murray JC. Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip. *Nat Genet*. 2008; 40(11): 1341-1347.

103. Ray JG, Vermeulen MJ, Wyatt PR, Cole DE. Association between folic acid food fortification and congenital orofacial clefts. *J Pediatr*. 2003; 143(6): 805-807.

104. Riley BM, Mansilla MA, Ma J, Daack-Hirsch S, Maher BS, Raffensperger LM, Russo ET, Vieira AR, Dodé C, Mohammadi M, Marazita ML, Murray JC. Impaired FGF signaling contributes to cleft lip and palate. *Proc Natl Acad Sci U S A*. 2007b 13; 104(11): 4512-4517.

105. Riley BM, Schultz RE, Cooper ME, Goldstein-McHenry T, Daack-Hirsch S, Lee KT, Dragan E, Vieira AR, Lidral AC, Marazita ML, Murray JC. A genome-wide linkage scan for cleft lip and cleft palate identifies a novel locus on 8p11-23. *Am J Med Genet A*. 2007a 15; 143A(8): 846-852.

106. Robin NH, Moran RT, Warman M, Ala-Kokko L. Stickler Syndrome. Editors In: Pagon RA, Bird TD, Dolan CR, Stephens K, Adam MP. *Source GeneReviews™ [Internet]. Seattle (WA): University of Washington, Seattle.* 1993-2000, updated 2011.

107. Schliekelman P, Slatkin M. Multiplex relative risk and estimation of the number of loci underlying an inherited disease. *Am J Hum Genet*. 2002; 71(6): 1369 - 85.

108. Schutte BC, Murray JC. The many faces and factors of orofacial clefts. *Hum Mol Genet*. 1999; 8(10): 1853-1859.

109. Shahrukh Hashmi S, Gallaway MS, Waller DK, Langlois PH, Hecht JT. National Birth Defects Prevention Study. Maternal fever during early pregnancy and the risk of oral clefts. *Birth Defects Res A Clin Mol Teratol*. 2010; 288(3): 186-194.

110. Shi J, Song T, Jiao X, Qin C, Zhou J. Single-nucleotide polymorphisms (SNPs) of the IRF6 and TFAP2A in non-syndromic cleft lip with or without cleft palate (NSCLP) in a northern Chinese population. *Biochem Biophys Res Commun*. 2011; 15;410(4): 732-736.

111. Shi M, Christensen K, Weinberg CR, Romitti P, Bathum L, Lozada A, Morris RW, Lovett M, Murray JC. Orofacial cleft risk is increased with maternal smoking and specific detoxification- gene variants. *Am J Hum Genet*. 2007; 80(1): 76-90.

112. Shi M, Wehby GL, Murray JC. Review on genetic variants and maternal smoking in the etiology of oral clefts and other birth defects. *Birth Defects Res C Embryo Today*. 2008; 84(1): 16-29.

113. Shields ED, Bixler D, Fogh-Andersen P. Facial clefts in Danish twins. *Cleft Palate J*. 1979; 16(1): 1-6.

114. Sperber GH. Formation of the primary and secondary palate. In: Wyszynski DF, ed. Cleft lip and palate: from origin to treatment. New York: *Oxford University Press*, 2002; 5-24.

115. Stanier P, Moore GE: Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. *Hum Mol Gen*. 2004; 13: 73-81.

116. Stavnezer E, Brodeur D, Brennan LA. The v-SKI oncogene encodes a truncated set of c-SKI coding exons with limited sequence and structural relatedness to v-myc. *Mol Cell Biol*. 1989; 9(9): 4038-4045.

117. Stein J, Mulliken JB, Stal S, Gasser DL, Malcolm S, Winter R, Blanton SH, Amos C, Seemanova E, Hecht JT. Nonsyndromic cleft lip with or without cleft palate: evidence of linkage to BCL3 in 17 multigenerational families. *Am J Hum Genet*. 1995; 57(2): 257-272.

118. Suazo J, Santos JL, Jara L, Blanco R. Association between bone morphogenetic protein 4 gene polymorphisms with nonsyndromic cleft lip with or without cleft palate in a chilean population. *DNA Cell Biol*. 2010; 29(2): 59-64.

119. Sutrave P, Hughes SH. Isolation and characterization of three distinct cDNAs for the chicken c-SKI gene. *Mol Cell Biol.* 1989; 9(9): 4046-4051.

120. Suzuki S, Marazita ML, Cooper ME, Miwa N, Hing A, Jugessur A, Natsume N, Shimozato K, Ohbayashi N, Suzuki Y, Niimi T, Minami K, Yamamoto M, Altannamar TJ, Erkhembaatar T, Furukawa H, Daack-Hirsch S, L'heureux J, Brandon CA, Weinberg SM, Neiswanger K, Deleyiannis FW, de Salamanca JE, Vieira AR, Lidral AC, Martin JF, Murray JC. Mutations in BMP4 are associated with subepithelial, microform, and overt cleft lip. *Am J Hum Genet.* 2009; 84(3): 406-411.

121. Thisse B, Thisse C. Functions and regulations of fibroblast growth factor signaling during embryonic development. *Dev Biol.* 2005; 287(2): 390-402.

122. Trarbach EB,Silveira LG,Latronico AC. Genetic insights into human isolated gonadotropin deficiency. *Pituitary.* 2007; 10(4): 381-391.

123. Trasler DG. Pathogenesis of cleft lip and its relation to embryonic face shape in A-J and C57BL mice. *Teratology.* 1968; 1(1): 33-49.

124. Trumpp A, Depew MJ, Rubenstein JL, Bishop JM, Martin GR. Cre-mediated gene inactivation demonstrates that FGF8 is required for cell survival and patterning of the first branchial arch. *Genes Dev.* 1999; 1;13(23): 3136-3148.

125. Turner N, Grose R. Fibroblast growth factor signalling: from development to cancer. *Nat Rev Cancer.* 2010; 10(2): 116-129.

126. van Rooij A, Wegerif MJ, Roelofs HM, Peters WH, Kuijpers-Jagtman AM, Zielhuis GA, Merkus HM, Steegers-Theunissen RP. Smoking, genetic polymorphisms in biotransformation enzymes, and nonsyndromic oral clefting: a gene-environment interaction. *Epidemiology.* 2001; 12(5): 502-507.

127. Vieira AR, Avila JR, Daack-Hirsch S, Dragan E, Félix TM, Rahimov F, Harrington J, Schultz RR, Watanabe Y, Johnson M, Fang J, O'Brien SE, Orioli IM, Castilla EE, Fitzpatrick DR, Jiang R, Marazita ML, Murray JC. Medical sequencing of candidate genes for nonsyndromic cleft lip and palate. *PLoS Genet.* 2005; 1(6): e64.

128. Wan M, Cao X. BMP signaling in skeletal development. *Biochem Biophys Res Commun.* 2005; 328: 651-657.

129. Wang H, Zhang T, Wu T, Hetmanski JB, Ruczinski I, Schwender H, Murray T, Fallin MD, Redett RJ, Raymond GV, Jin SC, Wu-Chou YH, Chen PK, Yeow V, Chong SS, Cheah FS, Jee SH, Jabs EW, Liang KY, Scott A, Beaty TH. The

FGF&FGFR Gene Family and Risk of Cleft Lip with/without Cleft Palate. *Cleft Palate Craniofac J.* 2011; [Epub ahead of print].

130. Warner DR, Pisano MM, Greene RM. Nuclear convergence of the TGFbeta and cAMP signal transduction pathways in murine embryonic palate mesenchymal cells. *Cell Signal.* 2003; 15(2): 235-242.

131. Warrington A, Vieira AR, Christensen K et al. Genetic evidence for the role of loci at 19q13 in cleft lip and palate. *J Med Genet.* 2006; 43(6): e26.

132. Wehby G, Cassell CH. The impact of orofacial clefts on quality of life and healthcare use and costs. *Oral Dis.* 2010; 16(1): 3-10.

133. Wehby GL, Murray JC. Folic acid and orofacial clefts: a review of the evidence. *Oral Dis.* 2010; 16(1) : 11-19.

134. Wilcox A, Lie RT, Solvoll K, Taylor J, McConnaughey DR, Abyholm F, Vindenes H, Vollset SE, Drevon CA. Folic acid supplements and risk of facial clefts: national population based case- control study. *BMJ.* 2007; 334(7591): 464.

135. Wu T, Liang KY, Hetmanski JB, Ruczinski I, Fallin MD, Ingersoll RG, Wang H, Huang S, Ye X, Wu-Chou YH, Chen PK, Jabs EW, Shi B, Redett R, Scott AF, Beaty TH. Evidence of gene- environment interaction for the IRF6 gene and maternal multivitamin supplementation in controlling the risk of cleft lip with/without cleft palate. *Hum Genet.* 2010; 128(4): 401-410.

136. Wyszynski DF, Maestri N, McIntosh I, Smith EA, Lewanda AF, Garcia-Delgado C, Vinageras-Guarneros E, Wulfsberg E, Beaty TH. Evidence for an association between markers on chromosome 19q and non-syndromic cleft lip with or without cleft palate in two groups of multiplex families. *Hum Genet.* 1997; 99(1): 22-26.

137. Yazdy M, Honein MA, Xing J. Reduction in orofacial clefts following folic acid fortification of the U.S. grain supply. *Birth Defects Res A Clin Mol Teratol.* 2007; 79(1): 16-23.

138. Zhu H, Kartiko S, Finnell RH. Importance of gene-environment interactions in the etiology of selected birth defects. *Clin Genet.* 2009; 75(5): 409-423.

139. Zucchero TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, Natsume N, Yoshiura K, Vieira AR, Orioli IM, Castilla EE, Moreno L, Arcos-Burgos M, Lidral AC, Field LL, Liu YE, Ray A, Goldstein TH, Schultz RE, Shi M, Johnson MK, Kondo S,

Schutte BC, Marazita ML, Murray JC. Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *N Engl J Med*. 2004; 19; 351(8): 769 - 80.

Appendix 1

List of LASHAL codes

Code	Description
.....	No Condition
.....*	Microform of cleft lip on left side
.....l	Partial left lip cleft
.....L	Complete left lip cleft
.....*l	Microform of alveolus cleft with partial lip cleft on the left side
.....al	Partial left lip-alveolus cleft
.....AL	Complete left lip-alveolus cleft
...*...	Bifid uvula
...s...	Incomplete soft palate cleft
...S...	Complete soft palate cleft
...S..L	Complete cleft of velum and left lip
...SHAL	Complete left unilateral cleft lip, alveolus, hard and sof palate cleft (CLAP)
..*s*..	Submucous velar cleft
..h.h..	Submucous clefts of hard palate only
..h*h..	Submucous cleft of hard palate c/ bifid uvula
..hsh..	Submucous clefts of hard and soft palate
..hShal	Partial left lip and alveolus cleft associated with a complete velar cleft extending incompletely into the hard palate
..HSH..	Complete cleft of hard and soft palate
*.....	Microform of cleft lip on right side
*******	Midline cleft
l.....	Partial cleft lip of the right side
l.....l	Partial bilateral cleft lip
l*.....	Partial right cleft lip with microform of right alveolus left
l*...*l	Bilateral microform of lip clefts
la.....	Partial right lip-alveolus cleft
la...al	Partial bilateral lip-alveolus cleft

Code	Description
lahSh..	Partial right lip and alveolus cleft associated with a complete velar cleft extending incompletely into the hard palate
lA.S.Al	Partial bilateral cleft lip and complete bilateral alveolus cleft associated with a complete soft palate cleft
L.....L	Complete bilateral lip cleft
L..S...	Complete cleft of velum and right lip
LahSHAL	Complete right lip and partial alveolus cleft associated with a complete left unilateral CLAP extending into the hard palate on the right side
LA...AL	Complete bilateral lip-alveolus cleft
LAHS...	Complete right unilateral CLAP cleft
LAHShal	Complete left lip and partial alveolus cleft associated with a complete right unilateral CLAP extending into the hard palate on the right side
..hSh..	Bilateral incomplete cleft hard palate with complete soft palate cleft
lahshal	palate/uvula
L*****L	Bilateral incomplete cleft lip (CL)
LAHS..L	Bilateral complete cleft lip with unilateral cleft palate
HSH	Unknown syn.cardiac murmur
lahs...	Incomplete cleft lip and palate right side
..hshaL	Unilateral CL left and partial cleft alveolus bilat and bifid uvula
laHSHAL	L-UCLP with R incomplete CL and aveolus Bilateral complete CP
lahS...	R incomplete CL, cleft palate
...SHaL	Complete cleft of hard and soft palate. cleft of left lip and incomplete cleft of alveolus

Code	Description
.AHSHAl	Left incomplete CL with complete cleft hard and soft palate
LAH....	R-cleft lip/alveolus/hard palate
LA.S.AL	Bilateral cleft lip and aveolus and bilateral cleft soft palate and uvula/hard palate intact
..*.*..	palatal fistula
.AHSHA.	cleft of alveolus&palate
.*HSH*.	Bilateral cleft hard and soft palate with bilateral microform cleft aveolus
S	2 degree palate only
***SHAL	Unilateral cleft lip and palate (left)
La.....	Complete cleft of left lip and incomplete cleft of left alveolus
..HsH.l	Bilateral complete hard bilateral incomplete soft microform lip - left side
L.....	Complete right lip cleft
L..S.AL	Complete left lip and alveolus cleft associated with a complete right cleft lip and a complete velar cleft
L.HSH.L	Complete bilateral cleft lip associated with a complete cleft of the hard and soft palate
LA.....	Complete right lip-alveolus cleft
LA.S..L	Complete right lip and alveolus cleft associated with a complete left cleft lip and complete velar cleft
LA.SHAL	Complete right lip and alveolus cleft associated with a complete unilateral left cleft lip, alveolus and palate
LAHS.AL	Complete left lip and alveolus cleft associated with a complete unilateral right cleft lip, alveolus and palate
LAHSHAL	Complete bilateral CLAP cleft
..***..	Unknown Cleft Diagnosis
L..SHAL	BCL w/ L-UCP L—SHAL

Code	Description
.....aL	Cleft lip left side with partial cleft of left alveolus
La...aL	Bilateral cleft lip and bilateral partial clefts of the alveolus
LAHSHAl	Left incomplete cleft lip with right complete CL and bilateral complete cleft palate
L....aL	bilateral cleft lip with left alveolar notch
....h*L	L-CL w/ L-incomplete cleft of hard palate with associated nasal deformity *aveolar notch
.*****.	Psuedocleft
..H.H..	Cleft of hard palate
SHAL	complete unilateral CLAP cleft
laal	Partial bilateral lip alveolar cleft
L.H.H.L	BCL with severe nasal deformity no cleft palate
l**s**l	Bilateral incomplete cleft lip and secondry cleft palate
l..S..l	bilateral incomplete cleft lip and soft palate
LASH..L	Bilateral cleft lip /unilateral cleft palate
....HAL	L-UCL with aveolar cleft
AL	Unilateral Cleft Lip with Alveous
...S.AL	Cleft soft palate with unilateral cleft lip
...S..	Soft Palate Cleft
LAHS	Unilateral Cleft Lip and Palate (right)
LA	Unilateral Cleft Lip& Alveolus
...sh.L	Unilateral incomplete cleft palate with unilateral complete cleft lip
LAH.HAL	Bilateral complete cleft lip & hard palate only
LA*.*..	Complete unilateral cleft lip(right)with partial cleft palate
LAHSH..	R UCLP
*....AL	Microform of cleft lip (RT)with complete cleft lip and alveolus

Code	Description
L.H....	Unilateral complete cleft lip(RT)with complete cleft of hard palate
*..SHAl	microform of cleft lip right side,partial cleft lip left side and complete cleft of alveolus and palate left side
...*.aL	Unilateral CL (Lt side) bifid uvula, cleft partially of alveolus lt
L..*..L	Bilateral cleft lip & bifid uvula
l..*.al	
...shal	Left incomplete lip
LahS...	Compelete right side cleft lip and soft palate with part right alveolus and palate
l..SHAl	Incomplete bilateral cleft lip and unilateral complete cleft palate
lAHS..l	Bilateral incomplete cleft lip, right side complete cleft palate
l.HSH.L	Bilateral complete cleft palate, complete rt CL and incomplete lt CL
lahSHAL	Complete LT CLP, incomplete RT CLP
...SHAl	Incomplete CL and complete CP left
LaHS...	Complete unilateral cleft lip/palate Rt side and incomplete cleft of alveolus
SH.L	Complete unilateral CLP left sided, no alveolar cleft
...SH.L	Unilateral complete CLP left sided, no alveolar cleft
LA...*L	Complete bilateral CL, complete cleft of rt alveolus and notched left alveolus
lAHSHAL	Incomplete rt CL, complete CP and lt CL
l..SHAL	Unilateral CLP lt side, incomplete CL rt side
laHSHal	Bilateral incomplete CL and alveolus, complete CP
l*....Al	Incomplete CL lt, complete lt alveolus, microform rt lip
L....AL	Complete bilateral CL, complete lt alveolus

Code	Description
LAHSh..	Complete CLP rt side, incomplete CP lt side
...S.al	Unilateral left CL and alveolus, complete soft palate cleft
LAHS..*	Complete CLAP rt side, left microform lip
..HSHAL	Complete bilateral CP, unilateral complete CL and alveolus lt
LAHSHal	Bilateral complete CP, complete rt CL, incomplete lt CL
..HSH.*	Microform notched lip Lt side with complete cleft of secondary palate
LAhS...	Unilateral complelte cleft of lip/alveolus Rt side with incomplete cleft of secondary palate
L*.....	Complete unilateral Rt cleft lip with notching of the aveolar arch

Appendix 2

Selected SNPs in *BCL3* gene for MALDI-TOF genotyping, PCR primer sequences and PCR fragment sizes used in the study

SNP^	PCR primers/ Minisequencing (MS) primers	PCR fragment size
rs7257231	F-5'CAGAGCATAGGGTCACCAG3' R-5'TCCCAAGGCACAGCTTAC3'	148 bp
rs10401176	F-5'AGCGTGACAGCTGGAGAG3' R-5'CAAATCCATACCAACCCAT3'	131 bp
rs8103315	F-5'GCACCCAGCAATTCATCA3' R-5'GCAGCTTCCTCTCCCTCTA3'	173 bp
rs2927457	F-5'TGAGACTTTACCGGAACG3' R-5'GCCTGTGAGGAGATGGAA3'	166 bp
rs11671085	F-5'GCCCAGCAGACCTGTTAC3' R-5'GCGAATGATTTCAGAGAAAC3'	137 bp
rs1979377	F-5'GTCCTCACCTCCCTTTTAGT3' R-5'GCAGTGGTGCTATCTTGTG3'	290 bp
rs2927456	F-5'TGAGGAATAAGGGTTCAGAA3' R-5'AATGTGGTGATCACAGCC3'	118 bp
rs2306148	F-5'GTCCAGCTCCGGTTAATT3' R-5'GAGCTGCCGGAGTACATT3'	236 bp

^ SNP - single nucleotide polymorphism

Appendix 3

PCR programs used for *BCL3* gene analysis with MALDI-TOF technology

SNP ID	Temperature	Time	Cycles
rs7257231	95°C	10 minutes	1x
	95°C	30 seconds	40x
	60°C	30 seconds	40x
	72°C	45 seconds	40x
	72°C	5 minutes	1x
rs10401176	95°C	10 minutes	1x
	95°C	30 seconds	40x
	60°C	30 seconds	40x
	72°C	45 seconds	40x
	72°C	5 minutes	1x
rs8103315, rs2927456	95°C	5 minutes	1x
	95°C	30 seconds	40x
	54°C	30 seconds	40x
	72°C	30 seconds	40x
	72°C	5 minutes	1x
rs1979377	95°C	10 minutes	1x
	95°C	30 seconds	40x
	54°C	30 seconds	40x
	72°C	30 seconds	40x
	72°C	5 minutes	1x

Appendix 4

The best results of single-marker association analysis (P≤0.05) associated with non-syndromic CL/CLP and CP

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
CL/CLP									
1	*SKI*	rs16824948	2176080	C/T	0.382	0.088	$0.0013x10^{-14}$	6.376	4.039-10.07
1	*SKI*	rs262683	2145729	C/T	0.231	0.15	0.0161	1.7	1.101-2.625
1	*IRF6*	rs630065	208065285	T/C	0.311	0.193	0.0014	1.886	1.276-2.787
1	*LHX8*	rs941032	75395204	C/T	0.429	0.533	0.0163	0.659	0.468-0.927
2	*TGFA*	rs7605323	70637335	T/C	0.396	0.267	0.0013	1.805	1.258-2.589
2	*TGFA*	rs6743202	70625029	A/T	0.434	0.311	0.0033	1.698	1.191-2.421
2	*TGFA*	rs10489985	70626228	A/G	0.277	0.189	0.0156	1.65	1.097-2.483
2	*TGFA*	rs3771477	70614252	G/A	0.13	0.073	0.0258	1.893	1.073-3.342
2	*TGFA*	rs2215021	70649765	C/T	0.425	0.334	0.0304	1.469	1.036-2.083
2	*FN1*	rs7609476	216002261	A/C	0.552	0.464	0.0418	1.425	1.013-2.005
2	*FN1*	rs1250233	216015348	G/A	0.25	0.18	0.0439	1.523	1.01-2.297
2	*FN1*	rs2577289	215998392	T/A	0.267	0.196	0.0507	1.493	0.998-2.234
4	*MSX1*	rs12498543	4934360	C/T	0.086	0.142	0.0476	0.566	0.321-0.999
5	*FGF1*	rs34010	141961149	G/T	0.24	0.395	0.0002	0.485	0.3309-0.71
5	*FGF1*	rs2070715	142004440	C/T	0.325	0.419	0.0285	0.67	0.467-0.959
8	*FGFR1*	rs7829058	38451252	G/C	0.137	0.019	$0.0024x10^{-5}$	7.991	3.435-18.59

Continuation of table

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
8	*FGFR1*	rs2288696	38405382	G/A	0.137	0.21	0.0287	0.596	0.374-0.951
9	*FOXE1*	rs7860144	99666705	G/A	0.302	0.414	0.0072	0.611	0.426-0.876
9	*FOXE1*	rs874004	99661939	C/G	0.519	0.42	0.0217	1.491	1.059-2.099
9	*FOXE1*	rs973473	99660551	G/T	0.245	0.326	0.041	0.672	0.459-0.985
11	*MMP13*	rs11225485	102312941	A/G	0.038	0.094	0.0126	0.378	0.172-0.833
11	*MMP13*	rs7119194	102312606	C/T	0.019	0.058	0.0266	0.312	0.106-0.922
12	*COL2A1*	rs1541408	46634619	C/G	0.165	0.097	0.0156	1.847	1.117-3.055
16	*MMP2*	rs11865658	54062587	C/T	0.071	0.025	0.0074	3.017	1.296-7.022
16	*MMP2*	rs1347653	54062541	G/T	0.259	0.181	0.0252	1.59	1.057-2.391
16	*MMP2*	rs837535	54053553	A/G	0.236	0.164	0.0343	1.575	1.032-2.402
17	*WNT3*	rs11655598	42223260	C/G	0.307	0.069	$0.0053x10^{-4}$	5.925	3.593-9.772
17	*WNT3*	rs11653738	42242117	T/C	0.429	0.309	0.0038	1.679	1.181-2.386
17	*WNT9B*	rs9894638	42267900	T/C	0.198	0.274	0.0429	0.656	0.436-0.988
17	*WNT9B*	rs12150651	42288568	T/C	0.095	0.158	0.0434	0.575	0.335-0.989
17	*WNT9B*	rs4968282	42313936	A/G	0.189	0.262	0.0444	0.654	0.431-0.991
17	*RARA*	rs506728	35787106	G/C	0.179	0.12	0.0507	1.6	0.996-2.57
18	*OFC11*	rs2006747	46222356	C/G	0.514	0.398	0.0068	1.603	1.138-2.258
18	*OFC11*	rs328149	42424469	A/G	0.08	0.141	0.0299	0.532	0.298-0.947

Continuation of table

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
19	*BCL3*	rs4803750	49939467	A/G	0.052	0.10	0.0449	0.496	0.247-0.996
20	*MMP9*	rs6073991	44089519	A/G	0.257	0.184	0.0386	1.539	1.021-2.319
22	*TIMP3*	rs10483165	31517342	T/C	0.066	0.13	0.0167	0.474	0.254-0.883
22	*TIMP3*	rs9609643	31581059	G/A	0.014	0.044	0.0509	0.309	0.089-1.072
23	*TIMP1*	rs5906437	47340290	A/G	0.291	0.391	0.0411	0.639	0.416-0.984
CP									
1	*SKI*	rs16824948	2176080	C/T	0.397	0.088	0.0011×10^{-7}	6.777	3.577-12.84
1	*SKI*	rs262683	2145729	C/T	0.259	0.15	0.0401	1.972	1.022-3.807
1	*SKI*	rs12562937	2219338	C/T	0.052	0.146	0.0489	0.318	0.096-1.054
1	*IRF6*	rs17389541	208053795	A/G	0.466	0.244	0.0005	2.692	1.524-4.756
1	*IRF6*	rs9430018	208063165	G/T	0.586	0.403	0.0088	2.101	1.196-3.69
1	*LHX8*	rs17096272	75355512	G/A	0.345	0.224	0.0448	1.829	1.008-3.319
4	*FGF2*	rs11938826	123992064	C/G	0.276	0.144	0.0112	2.271	1.19-4.335
4	*FGF2*	rs308395	123966392	C/G	0.035	0.13	0.0357	0.239	0.057-1.014
4	*MSX1*	rs6832405	4934637	G/T	0.155	0.283	0.0403	0.465	0.22-0.98
5	*FGF1*	rs34016	141979150	G/A	0.172	0.066	0.006	2.934	1.322-6.512
5	*FGF1*	rs7722035	142036157	G/C	0.345	0.522	0.0122	0.482	0.27-0.86
5	*FGF1*	rs10064637	142023936	A/G	0.138	0.285	0.0187	0.402	0.184-0.878

Continuation of table

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
5	*FGF1*	rs9324891	142032705	G/T	0.122	0.243	0.0386	0.427	0.187-0.976
5	*FGF1*	rs17208908	142040162	G/A	0.155	0.282	0.0424	0.468	0.2220.988
8	*FGFR1*	rs7829058	38451252	G/C	0.207	0.019	$0.0002x10^{-6}$	13.16	4.93-35.1
12	*COL2A1*	rs2071358	46652716	G/T	0.069	0.188	0.0257	0.32	0.112-0.915
12	*COL2A1*	rs12300271	46660883	T/C	0.035	0.122	0.0487	0.258	0.061-1.095
12	*COL2A1*	rs12721428	46655160	C/T	0.017	0.094	0.0488	0.168	0.023-1.253
16	*CDH1*	rs7188750	67400396	G/A	0.31	0.158	0.0048	2.408	1.29-4.494
16	*CDH1*	rs3785076	67408284	A/G	0.069	0.02	0.0296	3.71	1.052-13.11
16	*MMP2*	rs11865658	54062587	C/T	0.086	0.025	0.0157	3.7	1.194-11.46
17	*WNT3*	rs11655598	42223260	C/G	0.414	0.07	$0.0039x10^{-11}$	9.46	4.879-18.34
17	*WNT3*	rs11653738	42242117	T/C	0.517	0.309	0.0019	2.39	1.365-4.192
17	*WNT3*	rs199494	42224229	A/G	0.276	0.461	0.0082	0.445	0.241-0.82
17	*WNT3*	rs111769	42227151	C/T	0.535	0.373	0.0195	1.931	1.105-3.374
17	*RARA*	rs2077464	35700090	A/G	0.172	0.073	0.0128	2.644	1.2-5.824
18	*OFC11*	rs2006747	46222356	C/G	0.621	0.398	0.0015	2.477	1.4-4.383
18	*OFC11*	rs953570	43988463	A/T	0.155	0.314	0.0136	0.402	0.191-0.846
19	*BCL3*	rs8103315	49946008	C/A	0.035	0.127	0.0396	0.245	0.058-1.04
20	*BMP2*	rs1980499	6694498	T/C	0.661	0.467	0.0069	2.224	1.232-4.014

Continuation of table

Chr*	Gene	SNP^	Location	Alleles#	MAF**		p-value	OR^^	95% CI##
					Cases	Controls			
20	*BMP2*	rs7270163	6699316	A/G	0.035	0.146	0.019	0.208	0.049-0.879
22	*TIMP3*	rs11287	31588777	C/T	0.052	0.011	0.0247	4.882	1.064-22.4
22	*TIMP3*	rs13054779	31522189	T/C	0.13	0.056	0.0398	2.532	1.016-6.31
23	*TBX22*	rs195294	79168053	T/C	0.122	0.309	0.013	0.311	0.118-0.817
23	*TBX22*	rs5913168	79182602	C/T	0.051	0.207	0.0193	0.207	0.049-0.881

* Chr - chromosome;
^ SNP - single nucleotide polymorphism;
# Major allele is listed first;
** MAF - minor allele frequency;
^^ OR - odds ratio;
95% CI - 95% confidence interval

Appendix 5

Case-control association analysis of *BCL3* gene haplotypes associated with CL/CLP in Latvian population

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs7257231	rs10401176	rs8103315	rs1979377	rs2927456	*	*	*
WIN1	1	1	2	1	2	0.101	0.034	0.0007
WIN1	2	2	2	2	2	0	0.033	0.0056
WIN1	1	1	1	1	1	0.592	0.692	0.0067
WIN1	1	1	1	2	1	0.042	0.017	0.0358
WIN1	1	1	1	2	2	0.024	0.011	0.1307
WIN1	2	1	1	1	1	0.154	0.119	0.1846
WIN1	1	2	1	2	2	0.018	0.026	0.5513
WIN1	1	2	1	1	1	0.069	0.063	0.7667
*	rs7257231	rs10401176	rs8103315	rs1979377	*	*	*	*
WIN1	1	1	2	1	*	0.099	0.039	0.0005
WIN1	2	2	2	2	*	0	0.035	0.0043
WIN1	1	1	1	2	*	0.067	0.027	0.0066
WIN1	1	1	1	1	*	0.586	0.674	0.0177
WIN1	2	1	1	1	*	0.152	0.118	0.183
WIN1	2	2	1	2	*	0.01	0.018	0.3935

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN1	1	2	1	2	*	0.02	0.03	0.4329
WIN1	1	2	1	1	*	0.067	0.061	0.7834
*	rs10401176	rs8103315	rs1979377	rs2927456	*	*	*	*
WIN2	1	2	1	1	*	0.096	0.039	0.001
WIN2	2	2	2	2	*	0	0.031	0.0077
WIN2	1	1	2	1	*	0.04	0.017	0.0418
WIN2	1	1	2	2	*	0.026	0.012	0.0902
WIN2	2	1	2	1	*	0.004	0.016	0.157
WIN2	1	1	1	1	*	0.741	0.785	0.1724
WIN2	2	1	2	2	*	0.026	0.033	0.5993
WIN2	2	1	1	1	*	0.068	0.07	0.8948
*	rs7257231	rs10401176	rs8103315	*	*	*	*	*
WIN1	1	1	2	*	*	0.098	0.038	0.0006
WIN1	2	2	2	*	*	0	0.036	0.0038
WIN1	2	1	1	*	*	0.15	0.116	0.1798
WIN1	2	2	1	*	*	0.013	0.028	0.1994
WIN1	1	1	1	*	*	0.652	0.691	0.2683
WIN1	1	2	1	*	*	0.087	0.091	0.8769

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs10401176	rs8103315	rs1979377	*	*	*	*	*
WIN2	1	2	1	*	*	0.096	0.039	0.0009
WIN2	1	1	2	*	*	0.067	0.027	0.0058
WIN2	2	2	2	*	*	0	0.033	0.0058
WIN2	1	1	1	*	*	0.74	0.784	0.1758
WIN2	2	1	2	*	*	0.029	0.049	0.209
WIN2	2	1	1	*	*	0.067	0.07	0.9092
*	rs8103315	rs1979377	rs2927456	*	*	*	*	*
WIN3	2	1	1	*	*	0.097	0.044	0.0036
WIN3	2	2	2	*	*	0.004	0.029	0.0259
WIN3	1	1	1	*	*	0.807	0.85	0.1366
WIN3	1	2	1	*	*	0.044	0.033	0.4554
WIN3	1	2	2	*	*	0.049	0.044	0.773
*	rs7257231	rs10401176				*	*	*
WIN1	2	2	*	*	*	0.017	0.065	0.005
WIN1	2	1	*	*	*	0.148	0.116	0.1983
WIN1	1	1	*	*	*	0.752	0.73	0.5239
WIN1	1	2	*	*	*	0.083	0.089	0.7779

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs10401176	rs8103315	*	*	*	*	*	*
WIN2	1	2	*	*	*	0.093	0.04	0.002
WIN2	2	2	*	*	*	0.007	0.035	0.025
WIN2	2	1	*	*	*	0.093	0.119	0.2845
WIN2	1	1	*	*	*	0.807	0.806	0.9807
*	rs8103315	rs1979377	*	*	*	*	*	*
WIN3	2	1	*	*	*	0.096	0.044	0.0031
WIN3	2	2	*	*	*	0.004	0.031	0.0183
WIN3	1	1	*	*	*	0.808	0.849	0.1461
WIN3	1	2	*	*	*	0.092	0.076	0.4466
*	rs1979377	rs2927456	*	*	*	*	*	*
WIN4	2	2	*	*	*	0.052	0.073	0.2821
WIN4	2	1	*	*	*	0.044	0.034	0.5162
WIN4	1	1	*	*	*	0.904	0.893	0.628

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4
^ SNP - single nucleotide polymorphism; * Empty cell

Appendix 6

Case-control association analysis of *BCL3* gene haplotypes associated with CP in Latvian population

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs7257231	rs10401176	rs8103315	rs1979377	rs2927456	*	*	*
WIN1	2	2	2	2	2	0	0.034	0.1453
WIN1	2	2	1	1	1	0	0.013	0.3626
WIN1	2	2	1	2	1	0	0.012	0.3857
WIN1	1	1	1	1	1	0.733	0.694	0.5259
WIN1	1	2	1	2	2	0.016	0.026	0.6201
WIN1	2	1	1	1	1	0.132	0.115	0.7092
WIN1	1	2	1	1	1	0.069	0.057	0.717
WIN1	1	1	1	2	1	0.019	0.014	0.7434
WIN1	1	1	2	1	1	0.033	0.035	0.9326
*	rs7257231	rs10401176	rs8103315	rs1979377	*	*	*	*
WIN1	2	2	2	2	*	0	0.035	0.1332
WIN1	2	2	1	2	*	0	0.017	0.3009
WIN1	2	2	1	1	*	0	0.013	0.3694
WIN1	1	1	1	1	*	0.724	0.677	0.4452
WIN1	2	1	1	1	*	0.129	0.113	0.6962

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN1	1	2	1	1	*	0.067	0.056	0.7401
WIN1	1	1	1	2	*	0.018	0.025	0.7554
WIN1	1	1	2	1	*	0.032	0.035	0.9194
WIN1	1	2	1	2	*	0.03	0.031	0.9906
*	rs10401176	rs8103315	rs1979377	rs2927456	*	*	*	*
WIN2	2	2	2	2	*	0	0.032	0.1532
WIN2	1	1	1	1	*	0.854	0.799	0.2995
WIN2	2	1	2	2	*	0.016	0.032	0.4705
WIN2	1	2	1	1	*	0.032	0.036	0.869
WIN2	1	1	2	1	*	0.017	0.015	0.8719
WIN2	1	2	1	2	*	0.015	0.017	0.905
WIN2	2	1	1	1	*	0.066	0.069	0.9346
*	rs7257231	rs10401176	rs8103315	*	*	*	*	*
WIN1	2	2	2	*	*	0	0.036	0.1266
WIN1	2	2	1	*	*	0	0.03	0.1705
WIN1	1	1	1	*	*	0.742	0.696	0.4539
WIN1	2	1	1	*	*	0.129	0.113	0.7088
WIN1	1	2	1	*	*	0.097	0.088	0.825
WIN1	1	1	2	*	*	0.032	0.036	0.8725

Continuation of table

aplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs10401176	rs8103315	rs1979377	*	*	*	*	*
WIN2	2	2	2	*	*	0	0.034	0.1408
WIN2	1	1	1	*	*	0.854	0.788	0.2265
WIN2	2	1	2	*	*	0.031	0.049	0.5274
WIN2	1	1	2	*	*	0.018	0.025	0.7154
WIN2	1	2	1	*	*	0.032	0.036	0.873
WIN2	2	1	1	*	*	0.066	0.068	0.9546
*	rs8103315	rs1979377	rs2927456	*	*	*	*	*
WIN3	2	2	2	*	*	0	0.031	0.1591
WIN3	1	1	1	*	*	0.919	0.854	0.1593
WIN3	1	2	2	*	*	0.016	0.042	0.3145
WIN3	2	1	1	*	*	0.032	0.041	0.7422
WIN3	1	2	1	*	*	0.032	0.031	0.9667
*	rs7257231	rs10401176	*	*	*	*	*	*
WIN1	2	2	*	*	*	0	0.068	0.0345
WIN1	1	1	*	*	*	0.774	0.733	0.4791
WIN1	2	1	*	*	*	0.129	0.113	0.7061
WIN1	1	2	*	*	*	0.097	0.087	0.7863
*	rs10401176	rs8103315	*	*	*	*	*	*

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN2	2	2	*	*	*	0	0.038	0.1192
WIN2	1	1	*	*	*	0.871	0.809	0.2296
WIN2	1	2	*	*	*	0.097	0.116	0.6433
WIN2	2	2	*	*	*	0.032	0.037	0.8506
*	rs8103315	rs1979377	*	*	*	*	*	*
WIN3	2	1	*	*	*	0	0.034	0.1386
WIN3	1	1	*	*	*	0.919	0.852	0.147
WIN3	1	2	*	*	*	0.048	0.073	0.4674
WIN3	2	1	*	*	*	0.032	0.0405	0.7511
*	rs1979377	rs2927456	*	*	*	*	*	*
WIN4	2	2	*	*	*	0.016	0.073	0.0886
WIN4	1	1	*	*	*	0.952	0.893	0.1426
WIN4	2	1	*	*	*	0.032	0.034	0.9336

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4
^ SNP - single nucleotide polymorphism; * Empty cell

Appendix 7

Case-control association analysis of haplotypes associated with CL/CLP in *BCL3* gene in Brazilian population

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	Frequency		p-value
					Cases	Controls	
*	rs7257231	rs10401176	rs8103315	rs2927456	*	*	*
WIN1	1	2	1	2	0.046	0.023	0.0678
WIN1	2	1	1	1	0.255	0.287	0.2775
WIN1	1	2	1	1	0.076	0.093	0.3617
WIN1	1	1	2	1	0.108	0.094	0.4709
WIN1	2	1	1	2	0.02	0.014	0.5065
WIN1	1	1	1	2	0.042	0.039	0.8457
WIN1	1	1	1	1	0.454	0.451	0.9223
*	rs7257231	rs10401176	rs8103315	*	*	*	*
WIN1	2	1	1	*	0.274	0.3	0.3821
WIN1	1	1	2	*	0.109	0.094	0.4572
WIN1	1	2	1	*	0.121	0.115	0.7761
WIN1	1	1	1	*	0.496	0.491	0.8781
*	rs10401176	rs8103315	rs2927456	*	*	*	*
WIN2	2	1	2	*	0.049	0.025	0.0535
WIN2	1	1	1	*	0.706	0.736	0.3161
WIN2	2	1	1	*	0.075	0.093	0.3417

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	Frequency		p-value
					Cases	Controls	
WIN2	1	2	1	*	0.108	0.094	0.4693
WIN2	1	1	2	*	0.061	0.053	0.6111
*	rs7257231	rs10401176	*	*	*	*	*
WIN1	2	1	*	*	0.272	0.298	0.3801
WIN1	1	1	*	*	0.611	0.59	0.531
WIN1	1	2	*	*	0.118	0.112	0.7818
*	rs10401176	rs8103315	*	*	*	*	*
WIN2	1	1	*	*	0.766	0.791	0.3744
WIN2	1	2	*	*	0.108	0.094	0.5032
WIN2	2	1	*	*	0.126	0.115	0.6083
*	rs8103315	rs2927456	*	*	*	*	*
WIN3	1	1	*	*	0.779	0.827	0.0695
WIN3	1	2	*	*	0.113	0.079	0.0875
WIN3	2	1	*	*	0.109	0.094	0.4713

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - sliding window 3
^ SNP - single nucleotide polymorphism; * Empty cell

Appendix 8

***BCL3* gene case-control association analysis of haplotypes associated with CP in Brazilian population**

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	Frequency		p-value
					Cases	Controls	
*	rs7257231	rs10401176	rs8103315	rs2927456	*	*	*
WIN1	1	1	2	1	0.155	0.095	0.1072
WIN1	1	2	1	2	0.045	0.02	0.1741
WIN1	2	1	1	1	0.228	0.282	0.3205
WIN1	1	1	1	2	0.025	0.042	0.4763
WIN1	1	1	1	1	0.423	0.459	0.5511
WIN1	1	2	1	1	0.12	0.089	0.5612
WIN1	2	1	1	2	0.014	0.013	0.9537
*	rs7257231	rs10401176	rs8103315	*	*	*	*
WIN1	2	2	1	*	0.043	0.01	0.0321
WIN1	1	1	2	*	0.155	0.094	0.0988
WIN1	1	1	1	*	0.424	0.497	0.2222
WIN1	2	1	1	*	0.231	0.292	0.2621
WIN1	1	2	1	*	0.148	0.107	0.2865
*	rs10401176	rs8103315	rs2927456	*	*	*	*
WIN2	1	1	1	*	0.619	0.734	0.0357
WIN2	2	1	2	*	0.06	0.023	0.0723

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	Frequency		p-value
					Cases	Controls	
WIN2	1	2	1	*	0.155	0.094	0.1003
WIN2	2	1	1	*	0.131	0.094	0.3125
WIN2	1	1	2	*	0.036	0.055	0.4632
*	rs7257231	rs10401176	*	*	*	*	*
WIN1	2	2	*	*	0.034	0.012	0.1274
WIN1	1	2	*	*	0.156	0.106	0.1885
WIN1	2	1	*	*	0.239	0.291	0.3455
WIN1	1	1	*	*	0.57	0.592	0.7081
*	rs10401176	rs8103315	*	*	*	*	*
WIN2	1	1	*	*	0.655	0.791	0.0078
WIN2	2	1	*	*	0.191	0.115	0.0621
WIN2	1	2	*	*	0.155	0.094	0.1016
*	rs8103315	rs2927456	*	*	*	*	*
WIN3	1	1	*	*	0.75	0.827	0.1001
WIN3	2	1	*	*	0.155	0.094	0.1016
WIN3	1	2		*	0.095	0.079	0.6125

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - sliding window 3
^ SNP - single nucleotide polymorphism; * Empty cells

Appendix 9

Case-control association analysis of haplotypes associated with CL/CLP in 19q13 locus

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs35385129	rs10421283	rs2927438	rs419010	rs2075620	*	*	*
WIN1	1	1	2	1	1	0.058	2.903	0.0884
WIN1	2	1	1	1	1	0.022	1.231	0.2672
WIN1	2	2	1	1	1	0.047	1.102	0.2938
WIN1	1	2	1	1	2	0.032	0.887	0.3464
WIN1	1	2	1	2	2	0.038	0.735	0.3913
WIN1	1	1	2	2	2	0.012	0.589	0.443
WIN1	1	2	1	1	1	0.08	0.542	0.4615
WIN1	2	1	1	2	1	0.037	0.527	0.4679
WIN1	1	1	1	2	1	0.199	0.505	0.4773
WIN1	1	1	1	1	2	0.046	0.342	0.5584
WIN1	1	2	2	2	2	0.014	0.238	0.6255
WIN1	1	2	2	2	1	0.014	0.235	0.6278
WIN1	2	2	1	2	1	0.058	0.211	0.6461
WIN1	1	1	1	1	1	0.113	0.2	0.6552
WIN1	1	2	1	2	1	0.109	0.181	0.6705
WIN1	2	2	1	2	2	0.017	0.166	0.6836

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN1	1	1	2	2	1	0.031	0.061	0.8051
WIN1	1	1	1	2	2	0.042	0.061	0.8057
WIN1	2	1	1	2	2	0.017	0.029	0.8642
WIN1	1	2	2	1	1	0.014	0.021	0.8852
*	rs10421283	rs2927438	rs419010	rs2075620	rs875255	*	*	*
WIN2	1	2	1	1	1	0.056	0.031	0.1946
WIN2	1	1	1	1	2	0.025	0.05	0.1958
WIN2	2	1	2	1	2	0.032	0.059	0.2099
WIN2	1	1	1	2	2	0.026	0.045	0.3075
WIN2	1	2	1	1	2	0.047	0.028	0.3099
WIN2	1	1	2	1	1	0.221	0.184	0.3405
WIN2	1	2	2	2	2	0.021	0.011	0.3878
WIN2	2	1	1	1	2	0.027	0.042	0.4383
WIN2	2	1	1	1	1	0.064	0.083	0.4746
WIN2	2	1	2	2	2	0.071	0.055	0.5173
WIN2	2	1	1	2	2	0.034	0.044	0.6144
WIN2	2	2	2	2	2	0.022	0.016	0.644
WIN2	1	1	2	1	2	0.066	0.056	0.6618

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN2	1	1	2	2	2	0.055	0.059	0.8493
WIN2	1	1	1	1	1	0.086	0.091	0.8515
WIN2	1	2	2	1	1	0.036	0.033	0.858
WIN2	2	1	2	1	1	0.111	0.114	0.9314
*	rs35385129	rs10421283	rs2927438	rs419010	*	*	*	*
WIN1	1	1	2	1	*	0.10	0.06	0.1172
WIN1	1	2	1	1	*	0.077	0.12	0.2572
WIN1	2	1	1	1	*	0.008	0.02	0.3234
WIN1	1	1	1	1	*	0.132	0.162	0.3978
WIN1	1	1	2	2	*	0.062	0.046	0.4522
WIN1	2	2	1	1	*	0.039	0.054	0.4575
WIN1	1	1	1	2	*	0.254	0.23	0.5677
WIN1	2	1	1	2	*	0.067	0.056	0.6347
WIN1	2	2	1	2	*	0.069	0.074	0.826
WIN1	1	2	2	1	*	0.014	0.013	0.9419
WIN1	1	2	1	2	*	0.149	0.146	0.9487
WIN1	1	2	2	2	*	0.03	0.029	0.9537
*	rs10421283	rs2927438	rs419010	rs2075620	*	*	*	*

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN2	1	2	1	1	*	0.102	0.059	0.0877
WIN2	2	1	1	1	*	0.088	0.124	0.2435
WIN2	1	1	2	1	*	0.279	0.236	0.3045
WIN2	1	1	1	2	*	0.026	0.045	0.328
WIN2	1	1	1	1	*	0.105	0.134	0.3636
WIN2	2	1	2	2	*	0.077	0.057	0.3995
WIN2	1	2	2	2	*	0.021	0.011	0.4026
WIN2	2	1	2	1	*	0.139	0.164	0.4852
WIN2	2	1	1	2	*	0.032	0.041	0.6142
WIN2	2	2	2	2	*	0.02	0.015	0.6757
WIN2	2	2	2	1	*	0.012	0.015	0.7876
WIN2	2	2	1	1	*	0.012	0.013	0.8734
WIN2	1	1	2	2	*	0.053	0.056	0.8872
WIN2	1	2	2	1	*	0.033	0.031	0.8947
*	rs2927438	rs419010	rs2075620	rs875255	*	*	*	*
WIN3	1	1	1	2	*	0.051	0.086	0.1628
WIN3	1	1	2	2	*	0.06	0.087	0.2892

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN3	2	1	1	1	*	0.056	0.037	0.3253
WIN3	2	2	2	2	*	0.041	0.026	0.3623
WIN3	2	1	1	2	*	0.052	0.036	0.3909
WIN3	1	2	1	1	*	0.32	0.288	0.4783
WIN3	1	1	1	1	*	0.149	0.17	0.5462
WIN3	1	2	1	2	*	0.10	0.116	0.6004
WIN3	1	2	2	2	*	0.123	0.111	0.7047
WIN3	2	2	1	1	*	0.05	0.0444	0.7995
*	rs35385129	rs10421283	rs2927438	*	*	*	*	*
WIN1	1	1	2	*	*	0.163	0.105	0.0717
WIN1	1	2	1	*	*	0.223	0.254	0.455
WIN1	2	2	1	*	*	0.11	0.129	0.5587
WIN1	2	1	1	*	*	0.072	0.075	0.896
WIN1	1	1	1	*	*	0.389	0.394	0.9144
WIN1	1	2	2	*	*	0.043	0.043	1
*	rs10421283	rs2927438	rs419010	*	*	*	*	*
WIN2	1	2	1	*	*	0.10	0.06	0.119
WIN2	2	1	1	*	*	0.117	0.163	0.1794

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN2	1	1	1	*	*	0.14	0.182	0.2374
WIN2	1	2	2	*	*	0.062	0.045	0.4259
WIN2	1	1	2	*	*	0.322	0.286	0.4265
WIN2	2	1	2	*	*	0.216	0.22	0.9067
WIN2	2	2	2	*	*	0.03	0.029	0.9639
WIN2	2	2	1	*	*	0.014	0.013	0.9759
*	rs2927438	rs419010	rs2075620	*	*	*	*	*
WIN3	1	1	1	*	*	0.197	0.258	0.1436
WIN3	2	1	1	*	*	0.11	0.072	0.1668
WIN3	1	1	2	*	*	0.058	0.086	0.282
WIN3	2	2	2	*	*	0.039	0.025	0.3761
WIN3	1	2	2	*	*	0.131	0.114	0.5801
WIN3	1	2	1	*	*	0.414	0.40	0.7699
WIN3	2	2	1	*	*	0.05	0.046	0.8479
*	rs419010	rs2075620	rs875255	*	*	*	*	*
WIN4	1	2	2	*	*	0.064	0.089	0.3309
WIN4	2	2	2	*	*	0.167	0.14	0.4289
WIN4	1	1	2	*	*	0.104	0.126	0.4869

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN4	2	1	1	*	*	0.363	0.331	0.4913
WIN4	2	1	2	*	*	0.097	0.11	0.6637
WIN4	1	1	1	*	*	0.205	0.204	0.9777
*	rs35385129	rs10421283	*	*	*	*	*	*
WIN1	1	1	*	*	*	0.551	0.498	0.2719
WIN1	1	2	*	*	*	0.266	0.298	0.4742
WIN1	2	2	*	*	*	0.11	0.128	0.5651
WIN1	2	1	*	*	*	0.072	0.076	0.8861
*	rs10421283	rs2927438	*	*	*	*	*	*
WIN2	1	2	*	*	*	0.162	0.104	0.0677
WIN2	2	1	*	*	*	0.333	0.382	0.2946
WIN2	1	1	*	*	*	0.461	0.47	0.8506
WIN2	2	2	*	*	*	0.044	0.044	0.9717
*	rs2927438	rs419010	*	*	*	*	*	*
WIN3	1	1	*	*	*	0.262	0.345	0.0643
WIN3	2	1	*	*	*	0.109	0.074	0.2033
WIN3	2	2	*	*	*	0.097	0.07365	0.3846
WIN3	1	2	*	*	*	0.533	0.5073	0.6023

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs419010	rs2075620	*	*	*	*	*	*
WIN4	1	2	*	*	*	0.062	0.087	0.3378
WIN4	2	2	*	*	*	0.174	0.142	0.3723
WIN4	1	1	*	*	*	0.309	0.333	0.6028
WIN4	2	1	*	*	*	0.456	0.439	0.7213
*	rs2075620	rs875255	*	*	*	*	*	*
WIN5	1	2	*	*	*	0.201	0.236	0.3904
WIN5	1	1	*	*	*	0.568	0.535	0.4971
WIN5	2	2	*	*	*	0.231	0.229	0.963

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4; WIN5 - sliding window 5
^ SNP - single nucleotide polymorphism; * Empty cell

Appendix 10

Case-control association analysis of haplotypes associated with CP in 19q13 locus

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs35385129	rs10421283	rs2927438	rs419010	rs2075620	*	*	*
WIN1	1	1	1	1	2	0.108	0.047	0.0944
WIN1	1	2	1	1	2	0.07	0.039	0.3219
WIN1	1	1	2	2	1	0.011	0.033	0.3999
WIN1	1	2	2	2	1	0.002	0.016	0.4365
WIN1	2	1	1	1	1	0.052	0.031	0.4485
WIN1	1	1	1	2	1	0.159	0.204	0.476
WIN1	2	2	1	2	1	0.045	0.069	0.5295
WIN1	1	2	2	1	1	0.004	0.014	0.5709
WIN1	2	2	1	1	1	0.062	0.044	0.5811
WIN1	1	1	1	1	1	0.086	0.11	0.6223
WIN1	2	2	1	2	2	0.019	0.011	0.6369
WIN1	1	1	1	2	2	0.035	0.044	0.7663
WIN1	1	1	2	1	1	0.063	0.055	0.8132
WIN1	1	2	2	2	2	0.021	0.017	0.8174
WIN1	1	2	1	2	1	0.101	0.112	0.8201
WIN1	1	2	1	2	2	0.035	0.031	0.9012

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN1	2	1	2	2	1	0.028	0.026	0.9237
WIN1	1	2	1	1	1	0.078	0.076	0.9668
WIN1	2	1	1	2	2	0.021	0.022	0.9873
*	rs10421283	rs2927438	rs419010	rs2075620	rs875255	*	*	*
WIN2	1	1	1	2	2	0.103	0.048	0.1236
WIN2	2	1	1	2	2	0.096	0.048	0.1721
WIN2	2	1	1	1	2	0.071	0.039	0.3144
WIN2	2	1	2	1	2	0.034	0.069	0.3533
WIN2	1	2	2	1	1	0.014	0.039	0.3905
WIN2	1	1	2	1	1	0.134	0.18	0.4303
WIN2	2	2	2	1	1	0.004	0.012	0.6101
WIN2	2	1	1	1	1	0.055	0.073	0.6464
WIN2	2	1	2	2	2	0.052	0.043	0.7715
WIN2	1	1	2	1	2	0.036	0.045	0.7778
WIN2	1	2	1	1	1	0.031	0.024	0.7782
WIN2	2	1	2	1	1	0.11	0.119	0.8469
WIN2	1	1	1	1	2	0.049	0.056	0.8534
WIN2	1	2	1	1	2	0.032	0.028	0.8875

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN2	1	1	2	2	2	0.062	0.066	0.9235
WIN2	1	1	1	1	1	0.098	0.094	0.9274
WIN2	2	2	2	2	2	0.021	0.02	0.9304
*	rs35385129	rs10421283	rs2927438	rs419010	*	*	*	*
WIN1	1	1	2	2	*	0.016	0.047	0.3069
WIN1	2	2	1	1	*	0.086	0.0512	0.3263
WIN1	1	1	1	2	*	0.179	0.233	0.406
WIN1	2	1	1	1	*	0.045	0.025	0.4421
WIN1	1	1	2	1	*	0.08	0.057	0.5427
WIN1	1	1	1	1	*	0.192	0.161	0.5884
WIN1	1	2	2	2	*	0.018	0.032	0.5916
WIN1	2	2	1	2	*	0.06	0.076	0.6784
WIN1	1	2	1	1	*	0.131	0.113	0.7087
WIN1	1	2	2	1	*	0.007	0.012	0.7732
WIN1	1	2	1	2	*	0.138	0.142	0.9424
WIN1	2	1	1	2	*	0.05	0.052	0.9577
*	rs10421283	rs2927438	rs419010	rs2075620	*	*	*	*
WIN2	1	1	1	2	*	0.114	0.05	0.0792

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN2	2	1	1	2	*	0.083	0.043	0.2283
WIN2	1	2	2	1	*	0.008	0.035	0.3173
WIN2	1	1	2	1	*	0.179	0.231	0.4222
WIN2	2	2	2	1	*	0.003	0.016	0.4822
WIN2	2	1	2	1	*	0.139	0.173	0.5508
WIN2	2	2	1	1	*	0.007	0.014	0.6865
WIN2	2	1	2	2	*	0.06	0.047	0.6956
WIN2	1	2	1	1	*	0.067	0.054	0.712
WIN2	1	1	2	2	*	0.052	0.061	0.8046
WIN2	2	1	1	1	*	0.133	0.122	0.8254
WIN2	2	2	2	2	*	0.022	0.018	0.8452
WIN2	1	1	1	1	*	0.134	0.138	0.9366
*	rs2927438	rs419010	rs2075620	rs875255	*	*	*	*
WIN3	1	1	2	2	*	0.20	0.094	0.0279
WIN3	1	2	1	2	*	0.072	0.113	0.3943
WIN3	2	2	1	1	*	0.022	0.049	0.3944
WIN3	1	2	1	1	*	0.244	0.293	0.4824
WIN3	1	1	1	2	*	0.118	0.091	0.551

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN3	2	2	2	2	*	0.02	0.025	0.8171
WIN3	1	1	1	1	*	0.15	0.162	0.833
WIN3	2	1	1	1	*	0.031	0.034	0.9018
WIN3	1	2	2	2	*	0.11	0.105	0.9171
WIN3	2	1	1	2	*	0.034	0.034	0.9909
*	rs35385129	rs10421283	rs2927438	*	*	*	*	*
WIN1	1	2	2	*	*	0.028	0.046	0.575
WIN1	2	1	1	*	*	0.0971	0.077	0.6282
WIN1	1	1	1	*	*	0.3713	0.395	0.7515
WIN1	2	2	1	*	*	0.143	0.127	0.7616
WIN1	1	2	1	*	*	0.269	0.253	0.8131
WIN1	1	1	2	*	*	0.092	0.102	0.8217
*	rs10421283	rs2927438	rs419010	*	*	*	*	*
WIN2	1	2	2	*	*	0.016	0.046	0.3284
WIN2	1	1	1	*	*	0.242	0.187	0.3674
WIN2	2	1	1	*	*	0.214	0.163	0.376
WIN2	1	1	2	*	*	0.229	0.285	0.4143
WIN2	2	2	2	*	*	0.019	0.033	0.6215

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN2	1	2	1	*	*	0.073	0.056	0.6317
WIN2	2	1	2	*	*	0.196	0.218	0.7268
WIN2	2	2	1	*	*	0.0113	0.014	0.9039
*	rs2927438	rs419010	rs2075620	*	*	*	*	*
WIN3	1	1	2	*	*	0.195	0.091	0.0305
WIN3	2	1	1	*	*	0.014	0.05	0.2601
WIN3	1	2	1	*	*	0.317	0.402	0.2606
WIN3	1	1	1	*	*	0.269	0.256	0.8529
WIN3	1	2	2	*	*	0.113	0.108	0.9226
WIN3	2	2	2	*	*	0.023	0.024	0.9416
WIN3	2	1	1	*	*	0.07	0.068	0.9578
*	rs419010	rs2075620	rs875255	*	*	*	*	*
WIN4	1	2	2	*	*	0.217	0.099	0.0161
WIN4	2	1	1	*	*	0.269	0.342	0.3097
WIN4	2	1	2	*	*	0.068	0.109	0.3861
WIN4	1	1	2	*	*	0.152	0.127	0.637
WIN4	1	1	1	*	*	0.171	0.193	0.7159
WIN4	2	2	2	*	*	0.123	0.13	0.8867

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs35385129	rs10421283	*	*	*	*	*	*
WIN1	2	1	*	*	*	0.099	0.078	0.6153
WIN1	1	1	*	*	*	0.461	0.496	0.6483
WIN1	2	2	*	*	*	0.141	0.126	0.7739
WIN1	1	2	*	*	*	0.299	0.30	0.9909
*	rs10421283	rs2927438	*	*	*	*	*	*
WIN2	2	2	*	*	*	0.035	0.048	0.6795
WIN2	2	1	*	*	*	0.405	0.378	0.7149
WIN2	1	2	*	*	*	0.085	0.10	0.7483
WIN2	1	1	*	*	*	0.475	0.474	0.9931
*	rs2927438	rs419010	*	*	*	*	*	*
WIN3	1	1	*	*	*	0.457	0.349	0.143
WIN3	1	2	*	*	*	0.423	0.503	0.2951
WIN3	2	2	*	*	*	0.037	0.078	0.3031
WIN3	2	1	*	*	*	0.083	0.07	0.7463
*	rs419010	rs2075620	*	*	*	*	*	*
WIN4	1	2	*	*	*	0.213	0.096	0.0156
WIN4	2	1	*	*	*	0.333	0.448	0.1314

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN4	2	2	*	*	*	0.127	0.133	0.9017
WIN4	1	1	*	*	*	0.327	0.323	0.9618
*	rs2075620	rs875255	*	*	*	*	*	*
WIN5	2	2	*	*	*	0.34	0.229	0.09223
WIN5	1	1	*	*	*	0.44	0.535	0.214
WIN5	1	2	*	*	*	0.22	0.236	0.8062

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4; WIN5 - sliding window 5
^ SNP - single nucleotide polymorphism; * Empty cell

Appendix 11

Case-control association analysis of haplotypes associated with CL/CLP in *IRF6* gene

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs4844880	rs2013162	rs861019	rs2073487	rs642961	*	*	*
WIN1	2	1	2	1	2	0.069	0.024	0.0204
WIN1	1	1	2	1	2	0.182	0.108	0.0263
WIN1	1	1	1	1	1	0.344	0.43	0.0721
WIN1	2	1	1	1	1	0.093	0.059	0.1801
WIN1	2	2	2	2	1	0.035	0.058	0.2805
WIN1	1	2	2	2	1	0.276	0.321	0.3245
*	rs2013162	rs861019	rs2073487	rs642961	rs658860	*	*	*
WIN2	1	1	1	1	2	0.006	0.184	0.0155×10^{-6}
WIN2	2	2	2	1	2	0	0.13	0.0114×10^{-4}
WIN2	1	2	1	2	2	0.252	0.136	0.002
WIN2	1	1	1	1	1	0.431	0.302	0.0055
WIN2	2	2	2	1	1	0.311	0.248	0.1476
*	rs4844880	rs2013162	rs861019	rs2073487	*	*	*	*
WIN1	1	1	2	1	*	0.1912	0.112	0.0194
WIN1	2	2	2	1	*	0.069	0.028	0.042
WIN1	1	1	1	1	*	0.34	0.423	0.0782

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN1	2	1	1	1	*	0.092	0.059	0.1859
WIN1	2	2	2	2	*	0.034	0.057	0.2814
WIN1	1	2	2	2	*	0.273	0.32	0.2964
*	rs2013162	rs861019	rs2073487	rs642961	*	*	*	*
WIN2	1	2	1	2	*	0.249	0.131	0.0014
WIN2	2	2	2	1	*	0.308	0.375	0.1492
WIN2	1	1	1	1	*	0.432	0.484	0.2821
WIN2	1	2	1	1	*	0.012	0.011	0.9034
*	rs861019	rs2073487	rs642961	rs658860	*	*	*	*
WIN3	1	1	1	2	*	0.006	0.185	0.0131×10^{-6}
WIN3	2	2	1	2	*	0	0.13	0.0113×10^{-4}
WIN3	2	1	2	2	*	0.25	0.136	0.0022
WIN3	1	1	1	1	*	0.429	0.301	0.0061
WIN3	2	2	1	1	*	0.316	0.249	0.1266
*	rs4844880	rs2013162	rs861019	*	*	*	*	*
WIN1	1	1	2	*	*	0.197	0.112	0.0123
WIN1	2	1	2	*	*	0.068	0.028	0.0424
WIN1	1	1	1	*	*	0.338	0.424	0.0682

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN1	2	1	1	*	*	0.092	0.059	0.1836
WIN1	1	2	2	*	*	0.272	0.321	0.2653
WIN1	2	2	2	*	*	0.034	0.057	0.2851
*	rs2013162	rs861019	rs2073487	*	*	*	*	*
WIN2	1	2	1	*	*	0.26	0.141	0.00155
WIN2	2	2	2	*	*	0.308	0.377	0.1363
WIN2	1	1	1	*	*	0.432	0.482	0.2978
*	rs861019	rs2073487	rs642961	*	*	*	*	*
WIN3	2	1	2	*	*	0.247	0.132	0.001883
WIN3	2	2	1	*	*	0.312	0.373	0.1839
WIN3	1	1	1	*	*	0.429	0.48	0.2931
WIN3	2	1	1	*	*	0.012	0.014	0.8341
*	rs2073487	rs642961	rs658860	*	*	*	*	*
WIN4	1	1	2	*	*	0.006	0.187	0.0078×10^{-6}
WIN4	2	1	2	*	*	0	0.131	0.0085×10^{-4}
WIN4	1	2	2	*	*	0.247	0.133	0.0019
WIN4	1	1	1	*	*	0.435	0.302	0.004
WIN4	2	1	1	*	*	0.312	0.247	0.131

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs4844880	rs2013162	*	*	*	*	*	*
WIN1	2	1	*	*	*	0.156	0.086	0.0213
WIN1	1	2	*	*	*	0.268	0.32	0.2403
WIN1	2	2	*	*	*	0.038	0.058	0.3542
WIN1	1	1	*	*	*	0.538	0.537	0.977
*	rs2013162	rs861019	*	*	*	*	*	*
WIN2	1	2	*	*	*	0.265	0.14	0.0009
WIN2	2	2	*	*	*	0.306	0.378	0.1206
WIN2	1	1	*	*	*	0.429	0.483	0.2714
*	rs861019	rs2073487	*	*	*	*	*	*
WIN3	2	1	*	*	*	0.259	0.144	0.0023
WIN3	2	2	*	*	*	0.312	0.375	0.1688
WIN3	1	1	*	*	*	0.429	0.481	0.2884
*	rs2073487	rs642961	*	*	*	*	*	*
WIN4	1	2	*	*	*	0.247	0.133	0.0019
WIN4	2	1	*	*	*	0.312	0.378	0.1548
WIN4	1	1	*	*	*	0.441	0.49	0.3174
*	rs642961	rs658860	*	*	*	*	*	*

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN5	1	2	*	*	*	0.006	0.318	$0.0093x10^{-13}$
WIN5	1	1	*	*	*	0.747	0.549	$0.0244x10^{-3}$
WIN5	2	2	*	*	*	0.247	0.133	0.00193

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4; WIN5 - sliding window 5
^ SNP - single nucleotide polymorphism; * Empty cell

Appendix 12

Case-control association analysis of haplotypes associated with CP in *IRF6* gene

Haplotype	SNP^ 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
*	rs4844880	rs2013162	rs861019	rs2073487	rs642961	*	*	*
WIN1	2	1	2	1	1	0.055	0.007	0.0078
WIN1	1	1	2	1	2	0.027	0.109	0.0658
WIN1	1	1	1	1	1	0.308	0.438	0.0807
WIN1	2	1	2	1	2	0.05	0.023	0.254
WIN1	2	2	2	2	1	0.106	0.068	0.3326
WIN1	1	2	2	2	1	0.376	0.308	0.3336
WIN1	2	1	1	1	1	0.078	0.048	0.3736
*	rs2013162	rs861019	rs2073487	rs642961	rs658860	*	*	*
WIN2	1	1	1	1	2	0	0.20	0.0004
WIN2	1	2	1	1	1	0.058	0.004	0.001
WIN2	2	2	2	1	1	0.481	0.26	0.0013
WIN2	2	2	2	1	2	0	0.117	0.0093
WIN2	1	1	1	1	1	0.385	0.288	0.1641
WIN2	1	2	1	2	2	0.077	0.132	0.2697
*	rs4844880	rs2013162	rs861019	rs2073487	*	*	*	*
WIN1	2	1	2	1	*	0.092	0.03	0.0354

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN1	1	1	1	1	*	0.311	0.434	0.0984
WIN1	1	1	2	1	*	0.043	0.111	0.1318
WIN1	2	2	2	2	*	0.123	0.066	0.1501
WIN1	2	1	1	1	*	0.074	0.048	0.4556
WIN1	1	2	2	2	*	0.358	0.311	0.5039
*	rs2013162	rs861019	rs2073487	rs642961	*	*	*	*
WIN2	1	2	1	1	*	0.058	0.011	0.0186
WIN2	2	2	2	1	*	0.481	0.375	0.1491
WIN2	1	1	1	1	*	0.385	0.484	0.1864
WIN2	1	2	1	2	*	0.077	0.131	0.2764
*	rs861019	rs2073487	rs642961	rs658860	*	*	*	*
WIN3	1	1	1	2	*	0	0.199	0.0004
WIN3	2	2	1	1	*	0.481	0.259	0.0013
WIN3	2	1	1	1	*	0.058	0.007	0.0058
WIN3	2	2	1	2	*	0	0.117	0.0094
WIN3	1	1	1	1	*	0.385	0.287	0.1592
WIN3	2	1	2	2	*	0.077	0.131	0.2731
*	rs4844880	rs2013162	rs861019	*	*	*	*	*
WIN1	2	1	2	*	*	0.093	0.03	0.0336

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN1	1	1	1	*	*	0.311	0.434	0.097
WIN1	1	1	2	*	*	0.042	0.11	0.1326
WIN1	2	2	2	*	*	0.122	0.065	0.1474
WIN1	2	1	1	*	*	0.074	0.048	0.4482
WIN1	1	2	2	*	*	0.359	0.313	0.5129
*	rs2013162	rs861019	rs2073487	*	*	*	*	*
WIN2	2	2	2	*	*	0.481	0.377	0.158
WIN2	1	1	1	*	*	0.385	0.482	0.1939
WIN2	1	2	1	*	*	0.135	0.141	0.9052
*	rs861019	rs2073487	rs642961	*	*	*	*	*
WIN3	2	1	1	*	*	0.058	0.014	0.043
WIN3	2	2	1	*	*	0.481	0.373	0.1438
WIN3	1	1	1	*	*	0.385	0.481	0.1997
WIN3	2	1	2	*	*	0.077	0.132	0.2703
*	rs2073487	rs642961	rs658860	*	*	*	*	*
WIN4	1	1	2	*	*	0	0.196	0.0005
WIN4	2	1	1	*	*	0.481	0.255	0.001
WIN4	2	1	2	*	*	0	0.122	0.0077
WIN4	1	1	1	*	*	0.442	0.294	0.0342

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN4	1	2	2	*	*	0.077	0.133	0.2606
*	rs4844880	rs2013162	*	*	*	*	*	*
WIN1	1	1	*	*	*	0.363	0.546	0.015
WIN1	2	1	*	*	*	0.156	0.076	0.062
WIN1	2	2	*	*	*	0.132	0.067	0.1053
WIN1	1	2	*	*	*	0.349	0.311	0.5872
*	rs2013162	rs861019	*	*	*	*	*	*
WIN2	2	2	*	*	*	0.378	1.962	0.1614
WIN2	1	1	*	*	*	0.483	1.694	0.1931
WIN2	1	2	*	*	*	0.14	0.01	0.9199
*	rs861019	rs2073487	*	*	*	*	*	*
WIN3	2	2	*	*	*	0.481	0.375	0.1524
WIN3	1	1	*	*	*	0.385	0.481	0.2015
WIN3	2	1	*	*	*	0.135	0.144	0.8608
*	rs2073487	rs642961	*	*	*	*	*	*
WIN4	2	1	*	*	*	0.481	0.378	0.1614
WIN4	1	2	*	*	*	0.077	0.133	0.2606
WIN4	1	1	*	*	*	0.442	0.49	0.5309
*	rs642961	rs658860	*	*	*	*	*	*

Continuation of table

Haplotype	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	Frequency		p-value
						Cases	Controls	
WIN5	1	1	*	*	*	0.923	0.549	0.0378×10^{-5}
WIN5	1	2	*	*	*	0	0.318	0.0195×10^{-4}
WIN5	2	2	*	*	*	0.077	0.133	0.2606

WIN1 - sliding window 1; WIN2 - sliding window 2; WIN3 - sliding window 3; WIN4 - sliding window 4; WIN5 - sliding window 5
^SNP - single nucleotide polymorphism; * Empty cell