

**MOLECULAR CHARACTERIZATION OF CLEFT LIP WITH
OR WITHOUT CLEFT PALATE IN PAKISTAN**

**A THESIS SUBMITTED FOR THE PARTIAL FULFILLMENT OF
DEGREE OF DOCTOR OF PHILOSOPHY**

**BY
SAIRA MALIK**

**DEPARTMENT OF MICROBIOLOGY AND MOLECULAR GENETICS
AND SCHOOL OF BIOLOGICAL SCIENCES, UNIVERSITY OF THE
PUNJAB, LAHORE, PAKISTAN.**

OCTOBER 2011

Supervisor:

Sadaf Naz

DR. SADAF NAZ
Assistant Professor,
School of Biological Sciences,
University of the Punjab,
Lahore, Pakistan

Co-Supervisor:

DR. SHAHIDA HASNAIN
Director,
School of Biological Sciences,
University of the Punjab,
Lahore, Pakistan

Dedicated to the Prophet Muhammad (Peace be upon Him)

TABLE OF CONTENTS

	Page No.
LIST OF ABBREVIATIONS	i-iii
LIST OF TABLES	iv-v
LIST OF FIGURES	vi-viii
ACKNOWLEDGEMENTS	ix-x
ABSTRACT	xi-xii
CHAPTER 1 INTRODUCTION	1-26
Normal Lip and Palate development	4
Classification of Orofacial clefts	4-5
Prevalence	5-6
Pakistan's Geography, Demography, and the prevalence of CL/P	6-8
Problems Associated with CL/P	8
Causes of Non-syndromic CL/P	8-9
Risk of CL/P	9-11
<i>IRF6</i> and CL/P (Literature review)	11-22
Other genes in CL/P	22-25
CHAPTER 2 MATERIALS AND METHODS	27-55
IRB Approvals and Informed consent	27
History and Clinical Examination	27-28
Study subjects and the criteria of selection	28-30
Blood collection	30
Saliva collection	31
Genomic DNA extraction	31-32
DNA Dilutions	32
Primer Designing	32-35
Polymerase Chain Reaction (PCR)	35-43
Agarose Gel Electrophoresis	43

DNA Sequencing	43-44
T/A Cloning	44-46
Microsatellite Amplification and Haplotype analysis	46-48
Fast silver staining method	48
Analysis of VWS patients	48-49
Analysis of <i>IRF6</i> association study	49-55
Analysis of large consanguineous families	55
CHAPTER 3 RESULTS	56-126
Frequency of VWS in Pakistan	56
Novel mutations in <i>IRF6</i> gene	56-70
Known mutations identified in <i>IRF6</i> gene	70-78
Other VWS families with or without mutation	78-83
<i>IRF6</i> Association studies	83-93
<i>IRF6</i> Association studies in trios	93-100
Case-control studies	100-118
Analysis of large consanguineous families	118-126
CHAPTER 4 DISCUSSION	127-148
Phenotypes in VWS individuals	127-134
Frequency of VWS in Pakistan	134
Sex ratios of VWS in Pakistan	134
Mutations of <i>IRF6</i> in VWS individuals from Pakistan	135-141
<i>IRF6</i> Association studies	141-146
Non-syndromic CL/P in consanguineous Pakistani families	146-148
REFERENCES	149-163
APPENDIX	164-167
PUBLICATIONS	

LIST OF ABBREVIATIONS:

(NH ₄) ₂ SO ₄	Ammonium Sulphate
ABI	Applied Biosystems
AP-2 α	Activator protein-2
ARMS	Amplification refractory mutation system
BMPs	Bone Morphogenetic proteins
BSL-II	Biosafety level 2
CaCl ₂	Calcium chloride
cc	Cubic centimeter
CH ₃ COONa	Sodium acetate
CL	Cleft lip
CL/P	Cleft lip with or without cleft palate
CLAPP	Cleft lip and palate association of Pakistan
CLP	Cleft lip and cleft palate
cM	Centi-Morgan
CP	Cleft palate
DLX	Distalless
DNA BD	DNA binding domain
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide
EDTA	Ethylene-diamine tetra acetic acid
EGF	Epidermal growth factor
FGF	Fibroblast growth factor
FOXE1	Forkhead box protein E1
GABRB3	Gamma-aminobutyric acid receptor subunit β 3
GLI2	Zinc finger protein
GSC	Gooseoid
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IRB	Institution Review Board
IRF6	Interferon regulatory factor 6

KCl	Potassium chloride
LB	Luria-Bertani
LD	Linkage disequilibrium
LOD	Log of odds ratio
M	Molarity
μM	Micro molar
MgCl_2	Magnesium chloride
mM	Milli molar
MSX	Muscle segment
MTHFR	Methylenetetrahydrofolate reductase
Na_2EDTA	Disodium ethylene-diamine tetra acetic acid
NaCl	Sodium chloride
NaOH	Sodium hydroxide
nm	Nano meter
OFC	Orofacial clefts
OTX	Orthodontical
PBD	Protein binding domain
PCR	Polymerase chain reaction
PIC	Polymorphism information content
pmole	Pico mole
PPS	Popliteal pterygium syndrome
PVRL1	Poliovirus receptor-related 1
RARA	Retinoic acid receptor alpha
SATB2	Special AT-rich sequence binding protein 2
SDS	Sodium dodecyl sulfate
SHH	Sonic hedgehog
SKI	Sloan-Kettering Institute
SMIR	Smad-interferon regulatory factor-binding domain
SNP	Single nucleotide polymorphism
SPRY2	Sprouty homolog 2
SUMO1	Small ubiquitin-like modifier

T.E	Tris EDTA
T/A	Thymine/Adenine
<i>Taq</i>	<i>Thermus aquaticus</i>
TBX10	T-box developmental gene 10
TDT	Transmission disequilibrium test
TGF α	Transforming growth factor alpha
TGF β	Transforming growth factor beta
TP53	Tumor protein 53
TP73L	Tumor protein p73-like
UV	Ultraviolet
VWS	Van der Woude syndrome
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside

LIST OF TABLES

Table No.	Topic	Page No.
1.1	Genetic loci for orofacial clefts	10
2.1	Primer sequences for PCR and DNA sequencing of <i>IRF6</i>	33
2.2	Primer sequences for SNP genotyping of <i>IRF6</i> by tetra-primer ARMS-PCR	36
2.3	Primer sequences for mutational screening of normal controls by tetra-primer ARMS-PCR	37
2.4	Primer sequences for competitive allele specific PCR	38
2.5	Conditions for PCR amplification of different exons of <i>IRF6</i> gene	39
2.6	Conditions for tetra-primer ARMS-PCR for SNP genotyping	41
2.7	Conditions for mutational screening of normal controls by tetra-primer ARMS-PCR	42
2.8	Amplification conditions for sequencing PCR	45
2.9	PCR program for microsatellite markers amplification	47
2.10	Contingency table used in transmission disequilibrium test	52
3.1	Table of <i>IRF6</i> gene mutations in Pakistani VWS families	57
3.2	Table showing genotype frequencies and the minor allele frequencies of the four SNPs of <i>IRF6</i> in the normal controls	92
3.3	Table showing the results of transmission disequilibrium test in trios	94
3.4	Pair-wise linkage disequilibrium between the four markers in trios	96
3.5	LOD scores between different markers in the trio study calculated by Haploview program	98
3.6	Table showing haplotype associations between <i>rs2235375</i> and <i>rs2013162</i> in the trio study	101
3.7	Table showing haplotype associations between <i>rs2235375</i> , <i>rs2013162</i> , and <i>rs1319435</i> in the trio study	103
3.8	Table showing haplotype associations between <i>rs2235543</i> , <i>rs2235375</i> , <i>rs2013162</i> , and <i>rs1319435</i> in the trio study	104
3.9	Table showing the results of transmission disequilibrium test in the case-control study	106
3.10	Table showing the genotypes transmitted in cases and controls	107

Table No.	Topic	Page No.
3.11	Odds ratio for genotypes in case-control study	108
3.12	Pair-wise linkage disequilibrium between the four markers in the case-control study	110
3.13	LOD scores between different markers in the case-control study calculated by the Haploview program	111
3.14	Table showing haplotype associations between <i>rs2235375</i> and <i>rs2013162</i> in the case-control study	113
3.15	Table showing haplotype associations between <i>rs2013162</i> and <i>rs1319435</i> in the case-control study	115
3.16	Table showing haplotype associations between markers <i>rs2235375</i> , <i>rs2013162</i> , and <i>rs1319435</i> in the case-control study	117
3.17	Table showing haplotype associations between markers <i>rs2235543</i> , <i>rs2235375</i> , <i>rs2013162</i> , and <i>rs1319435</i> in the case-control study	120
3.18	Table showing the results from Merlin of the highest scoring markers for family CLP-SM2	125

LIST OF FIGURES

Figure No.	Topic	Page No.
1.1	Photographs of different phenotypes of cleft lip	2
1.2	Photographs of different types of cleft palate	3
1.3	Different ethnic groups in Pakistan	7
1.4	Diagrammatic representation of <i>IRF6</i>	12
1.5	Photographs showing VWS phenotype	14
3.1	Percentage of <i>IRF6</i> mutations in Pakistani VWS patients	58
3.2	Percentage of different mutations of <i>IRF6</i>	59
3.3	Percentage of novel and known mutations found in <i>IRF6</i>	60
3.4	Number and distribution of novel mutations in different exons	61
3.5	Family VWS-SM17 and identification of a novel mutation	63
3.6	Family VWS-SM20 and identification of a novel mutation	64
3.7	Family VWS-SM15 and identification of a novel mutation	65
3.8	Family VWS-SM2 and identification of a novel mutation	67
3.9	Family VWS-SM16 and identification of a novel mutation	68
3.10	Family VWS-SM19 and identification of a novel mutation	69
3.11	Family VWS-SM3 and identification of a novel mutation	71
3.12	Family VWS-SM9 and mutation identification	72
3.13	Family VWS-SM18 and mutation identification	74
3.14	Family VWS-SM12 and mutation identification	75
3.15	Family VWS-NU1 and mutation identification	76
3.16	Family VWS-SM6 and mutation identification	77
3.17	Families VWS-SM1, VWS-SM10 and VWS-SM21	79
3.18	Family VWS-SM5 and mutation identification	80
3.19	Family VWS-SM7 and mutation identification	81
3.20	Family VWS-SM13	82
3.21	Families VWS-SM8, VWS-SM11, and VWS-SM22	84
3.22	Family VWS-SM4	85

Figure No.	Topic	Page No.
3.23	Ratio of affected males to affected females in <i>IRF6</i> association study	87
3.24	Bar graph showing the statistics of isolated cleft lip in single affected individuals of <i>IRF6</i> association study	88
3.25	Bar graph showing the statistics of cleft lip with cleft palate in single affected individuals of the <i>IRF6</i> association study	89
3.26	Gel photographs of tetra-primer ARMS PCR for SNP genotyping of <i>rs2235543</i> and <i>rs2235375</i>	90
3.27	Gel photographs of tetra-primer ARMS PCR for SNP genotyping of <i>rs2013162</i> and <i>rs1319435</i>	91
3.28	LD plots for SNPs <i>rs2235543</i> , <i>rs2235375</i> , <i>rs2013162</i> and <i>rs1319435</i>	95
3.29	LD blocks for <i>rs2235375</i> and <i>rs2013162</i>	97
3.30	LD block and haplotypes between <i>rs2235375</i> , <i>rs2013162</i>	99
3.31	LD block showing <i>rs2235375</i> , <i>rs2013162</i> and <i>rs1319435</i>	102
3.32	LD plot showing the presence of linkage disequilibrium between <i>rs2235375</i> and <i>rs2013162</i> and <i>rs2013162</i> and <i>rs1319435</i> in the case-control study	109
3.33	LD block between <i>rs2235375</i> and <i>rs2013162</i> in the case-control study	112
3.34	LD block between <i>rs2013162</i> and <i>rs1319435</i> in the case-control study	114
3.35	LD block for markers <i>rs2235375</i> , <i>rs2013162</i> and <i>rs1319435</i> and the haplotypes in the case-control study	116
3.36	LD block for the markers <i>rs2235543</i> , <i>rs2235375</i> , <i>rs2013162</i> and <i>rs1319435</i> and their haplotypes in the case-control study	119
3.37	Pedigree of family CLP-SM1 segregating non-syndromic CL/P	121
3.38	Pedigree of family CLP-SM2 with isolated CL/P	122
3.39	Pedigree of family CLP-SM3	123
3.40	Pedigree of family CLP-SM2 with STR marker data from one region in chromosome 12	124
4.1	Photographs showing unusual phenotypes in family VWS-SM4	129

Figure No.	Topic	Page No.
4.2	Photographs showing less commonly reported and unusual lip pit phenotypes in some VWS families	132
4.3	Clustal W alignment of the carboxy end of pfam 10401 showing the conservation of IRF6 Ser 407 to IRF5 Ser 421	138

ACKNOWLEDGEMENTS

All praise to the Almighty Allah, Who always makes me feel His presence when I feel lost and disappointed. Allah Who always chose the best for me and made it possible for me to complete my research work and finally end up with this piece of writing.

This dissertation has the honor of passing the scrutiny of the ever-vigilant and highly scientific mind of my supervisor, Dr. Sadaf Naz. I feel lucky to have benefited from her caring supervision throughout the length and breadth of this research work. She always provided me with the best advice in all the matters of my life.

Credit, in no small measure, is also due to Professor Dr. Shahida Hassnain, my co-supervisor to whom I am greatly indebted. Her guidance has been like a lamplight I have studied and researched under. She always boosted me up, raised my morale and paved the way for me to walk, run and to finally reach my destination.

I am also really thankful to Professor Dr. Javed Iqbal who provided me the opportunity to work in the great working environment of the school of Biological sciences. I also express my gratitude to Dr. Naeem Rashid for his support and encouragement.

I owe sincere and earnest thanks to the plastic surgeons, Dr. Tajammul Hussain and Dr. Ghulam Qadir Fayyaz, who made access to the cleft patients possible.

I would also like to thank all the staff of the plastic surgery hospital in general and Mr. Ejaz Ahmed in particular who supported me and provided me the opportunity for timely patient contact and sample collection. I am also thankful to Mr. Zaidullah for his untiring efforts on the project.

I am truly indebted and thankful to all the families for their participation and cooperation.

My special thanks to the administration of the School of Biological Sciences and Department of Microbiology and Molecular Genetics for facilitating my work.

This research would not have been possible without the help and support of my lab fellows, colleagues and students.

I would also like to show my gratitude to my family who always encouraged me and provided me with moral and financial support.

It is a great pleasure for me to thank everyone who helped me write my dissertation successfully.

Saira Malik

ABSTRACT

Orofacial clefts including cleft lip and cleft palate are birth defects which may be present either alone as isolated or non-syndromic cleft, or they may be present in association with other anomalies as a syndrome. Van der Woude syndrome (VWS) is one of the most common syndromes which involves orofacial clefts and is usually one of the less severe syndromes of clefting. The most characteristic feature of VWS is the presence of lower lip pits which are present in about 85% of the cases. Mutations and polymorphisms in the gene Interferon regulatory factor 6 (*IRF6*) are known to cause both syndromic as well as non-syndromic forms of cleft lip and/or cleft palate in humans.

To date, no research has been performed on genetics of cleft lip with or without cleft palate in Pakistan. In the current study, 23 VWS patients were recruited from different areas of Punjab, Azad Kashmir and Balochistan. Mutational screening of *IRF6* gene was performed on all collected samples. An association study between single nucleotide polymorphisms of *IRF6* and non-syndromic cleft lip with or without cleft palate (CL/P) in single affected individuals was also performed. Additionally, 3 large consanguineous families with an apparent recessive mode of inheritance were recruited in search of novel loci for non-syndromic CL/P. The frequency of VWS identified in the current study was calculated to be 1.17% from the Punjab province and 1% from Azad Kashmir which was lower as compared to other world populations. Additionally 15 mutations including 7 novel and 8 known mutations within the coding region of *IRF6* were identified which cause VWS. These mutations were found in 18 of 23 unrelated families segregating VWS from Pakistan. The newly identified mutations include 3 missense mutations (p.M1T, p.G99S, and p.S407P) while the 4 novel frameshift mutations are p.R7fsX52, p.A190fsX34, p.S212fsX12, and p.209fsX15. The known pathogenic variants included 6 missense and 2 nonsense mutations including p.R84C, p.V321M, p.R250Q, p.C374W, p.R400W, and p.R400Q, p.Q204X, and p.R412X. The present findings add to the spectrum of *IRF6* mutations which are responsible for VWS. Additionally, the identification of the same mutations responsible for VWS in Pakistan, as reported in other populations worldwide, marks these residues as mutational hotspots.

In the association studies between single nucleotide polymorphisms in and around *IRF6* (*rs2235543*, *rs2235375*, *rs2013162*, and *rs1319435*) and non-syndromic CL/P in case-parent trios from Punjab, the transmission disequilibrium test (TDT) showed highly significant association of the 'C' allele of *rs2235375* with non-syndromic CL/P ($P=0.008$). These results identify *IRF6* as a significant contributor to non-syndromic CL/P in Pakistan.

Three large consanguineous families with non-syndromic CL/P and recessive mode of inheritance were also analyzed in search of novel loci. Linkage analysis suggests that regions on chromosome 9 are likely implicated in orofacial clefting in family CLP-SM1. Other regions, such as those on chromosome 5, may also be significant in the inheritance of CL/P in family CLP-SM1. Beyond the dubious region on chromosome 12, family CLP-SM2 did not provide any obvious candidates of linkage. Based on these findings, it is expected that there will be more complexity in the inheritance of CL/P than single allele acting through a classical recessive or dominant model.

The results of this study have demonstrated for the first time the important role of *IRF6* in clefting in the Pakistani population. It has also further revealed the genetic complexity of non-syndromic cleft lip with or without cleft palate even in families which seem to segregate the disorder in an apparent Mendelian inheritance pattern. Identification of responsible genes will reveal pathways necessary for normal orofacial development.

CHAPTER 1
INTRODUCTION