

Supplementary Materials

Tables S1–S7; Figures S1 and S2

A Comprehensive Genetic Analysis of Slovenian Families with Multiple Cases of Orofacial Clefts Reveals Novel Variants in the Genes *IRF6*, *GRHL3*, and *TBX22*

Lara Slavec^{1,2}, Ksenija Geršak^{1,3}, Andreja Eberlinč⁴, Tinka Hovnik^{5,6}, Luca Lovrečič^{3,7}, Irena Mlinarič-Raščan², Nataša Karas Kuželički²

¹ University Medical Centre Ljubljana, Division of Gynaecology and Obstetrics, Research Unit, Ljubljana, Slovenia

² University of Ljubljana, Faculty of Pharmacy, Department of Clinical Biochemistry, Ljubljana, Slovenia

³ University of Ljubljana, Faculty of Medicine, Department of Gynaecology and Obstetrics, Ljubljana, Slovenia

⁴ University Medical Centre Ljubljana, Department of Maxillofacial and Oral Surgery, Ljubljana, Slovenia

⁵ University Medical Centre Ljubljana, University Children's Hospital, Clinical Institute for Special Laboratory Diagnostics, Ljubljana, Slovenia

⁶ University of Ljubljana, Faculty of Medicine, Institute of Biochemistry and Molecular Genetics, Ljubljana, Slovenia

⁷ University Medical Centre Ljubljana, Clinical Institute of Genomic Medicine, Ljubljana, Slovenia

***Corresponding author:**

Assoc. prof. Nataša Karas Kuželički, PhD

Tel: +386-1-47-69-629

Email: natasa.karas@ffa.uni-lj.si

Table S1. The fold change in copy number ($\Delta K Ct$) of *TBX22* coding exons calculated for the samples of family 7 (F-7) and controls

<i>TBX22</i> exon number	Fold change in copy number ($\Delta K Ct$) ^a															
	2		3		4		5		6		7		8		9	
Reference gene	<i>G6PD</i>	<i>IRF6</i>	<i>G6PD</i>	<i>IRF6</i>	<i>G6PD</i>	<i>IRF6</i>	<i>G6PD</i>	<i>IRF6</i>	<i>G6PD</i>	<i>IRF6</i>	<i>G6PD</i>	<i>IRF6</i>	<i>G6PD</i>	<i>IRF6</i>	<i>G6PD</i>	<i>IRF6</i>
Sample (sex)																
Normal controls																
Control 1 (F)	-0.074	-0.114	-0.042	-0.080	0.006	0.090	0.221	0.184	0.002	-0.036	-0.006	0.076	-0.003	0.103	-0.093	0.063
Control 2 (F)	0.035	-0.017	-0.016	-0.066	-0.115	-0.042	0.002	-0.046	-0.095	-0.145	-0.125	-0.054	-0.159	-0.064	-0.164	-0.023
Control 3 (F)	0.107	-0.024	0.085	-0.040	-0.041	-0.048	0.126	0.004	-0.018	-0.143	-0.020	-0.023	-0.060	-0.040	-0.083	-0.040
Control 4 (F)	-0.068	0.154	-0.027	0.185	0.150	N.A.	-0.349	-0.141	0.111	0.323	0.151	N.A.	0.221	N.A.	0.341	N.A.
Control 5 (M)	0.152	0.184	-0.035	-0.006	-0.011	0.019	-0.025	0.004	0.038	0.070	-0.007	0.023	-0.017	0.014	-0.032	0.004
Control 6 (M)	-0.058	-0.066	0.045	0.038	-0.068	-0.076	0.165	0.157	-0.032	-0.040	-0.114	-0.121	-0.107	-0.115	0.033	0.024
Control 7 (M)	-0.094	-0.118	-0.010	-0.032	0.080	0.057	-0.140	-0.162	-0.007	-0.030	0.121	0.099	0.124	0.101	-0.001	-0.028
Test samples																
F-7 IV-2 (M)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
F-7 III-4 (F)	-0.678	-0.900	-0.841	-1.053	-0.851	-0.948	-0.855	-1.063	-0.963	-1.176	-0.828	-0.918	-0.884	-0.951	-0.654	-0.650
F-7 III-2 (F)	-0.702	-1.084	-0.835	-1.200	-0.787	-1.042	-0.790	-1.147	-0.944	-1.309	-0.805	-1.046	-0.855	-1.073	-0.740	-1.006
F-7 IV-1 (M)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

F, female; M, male; x, no presence of real-time quantitative PCR (qPCR) product after amplification of tested *TBX22* exon, but qPCR product present when amplifying the exon of the reference gene; N.A., data not available

^a, $\Delta K Ct$ values of 0 ± 0.35 indicate no copy number change (no genetic abnormality) in males and females, whereas -1 ± 0.35 indicates loss of one allelic copy (the deletion of the *TBX22* exon) in females. For the male samples, loss of a single allelic copy was indicated when no qPCR product was present

Table S2. Rare variants identified in families with non-syndromic orofacial clefts whose involvement could not be clearly determined based on the results of *in silico* prediction tools, co-segregation analyses, and the literature

Gene (Exon)	Genomic location (hg19)	Ref>Alt	Nucleotide change ^a 'rsID' number	Amino acid change	Variant type	ACMG class	GnomAD ^b	Predicted deleterious (Y/N) ^c	Franklin's aggregated prediction	ClinVar ^d	Segregation (Y/N) ^e
<i>DLG1</i> (20)	Chr3:196,793,628	ATTAGA AAGTGA AGGCAC AG>A	NM_001366207.1: c.2048-22_2048-4del rs758466083	NP_001353136.1: p.?	splice-site indel	VUS (PM2)	8.92E-06	na	na	na	x
<i>FGFR1</i> (4)	Chr8:38,285,913	G>GTCA	NM_023110.3: c.396_398dup rs138489552	NP_075598.2: p.Asp133dup	in-frame indel	VUS (PM2,PM4, BP6)	4.59E-04	na	na	conflict	Y & N
<i>BMP4</i> (3)	Chr14:54,418,669	G>C	NM_001202.6: c.272C>G rs121912767	NP_001193.2: p.Ser91Cys	missense	VUS (PM2,PP3, BP6)	3.86E-04	Y	Deleterious	conflict	Y
<i>JAG2</i> (24)	Chr14:105,611,347	T>C	NM_002226.5: c.3004A>G	NP_002217.3: p.Arg1002Gly	missense	VUS (PM2)	2.41E-05	N	Uncertain	na	Y & N
<i>TBX22</i> (9)	chrX:79,286,536	G>A	NM_001109878.2: c.1489G>A rs761750322	NP_001103348.1: p.Asp497Asn	missense	VUS (PM2)	2.23E-04	N	Uncertain	na	Y

Y/N, yes (Y) or no (N); na, not available; conflict, conflicting interpretations of pathogenicity (uncertain significance, likely benign or benign); x, phenotype co-segregation could not be determined because the sample from the proband's affected family member was not available for the analysis; Y & N, variant segregates with phenotype but is also present in unaffected proband's sibling (possible reduced penetrance); Ref>Alt, reference and alternative allele, where reference allele is listed first

^a, All listed variants were found in one family and are all heterozygous, except for rs761750322 (*TBX22*), which is hemizygous; ^b, The frequency of the alternative allele from the Genome Aggregation Database (GnomAD); ^c, Predicted deleterious by *in silico* tools – missense SNV predicted deleterious by at least three *in silico* prediction tools (SIFT, PolyPhen-2, MutationAssessor, MutationTaster, FATHMM, CADD, MetaLR, REVEL or GERP++); ^d, Interpretation in ClinVar database; ^e, Variant co-segregates with disease phenotype in the family

Table S3. Rare variants identified in families with non-syndromic orofacial clefts excluded after co-segregation analysis because they do not segregate with the disease phenotype

Gene (Exon)	Genomic location (hg19)	Ref>Alt	Nucleotide change ^a 'rsID' number	Amino acid change	Variant type	ACMG class	GnomAD ^b	Predicted deleterious (Y/N) ^c	Franklin's aggregated prediction	ClinVar ^d
<i>MTHFR</i> (7)	chr1:11,854,790	G>A	NM_005957.5: c.1162C>T rs200138092	NP_005948.3: p.Arg388Cys	missense	Likely Pathogenic (PM2,PM5, PM1,PP3,PP2)	3.36E-04	Y	Deleterious	uncertain significance
<i>MTHFR</i> [#] (6)	chr1:11,855,315	C>T	NM_005957.5: c.871G>A rs771684838	NP_005948.3: p.Asp291Asn	missense	VUS (PM2,PP3, PP2)	2.75E-05	Y	Deleterious	na
<i>TBX10</i> (3)	chr11:67,402,333	A>T	NM_005995.5: c.331T>A rs117056541	NP_005986.2: p.Tyr111Asn	missense	VUS (PP3,PM2, BS2)	0.004743371	Y	Deleterious	na
<i>MMP3</i> (4)	chr11:102,712,927	C>A	NM_002422.5: c.583G>T rs782241657	NP_002413.1: p.Ala195Ser	missense	VUS (PM2)	1.76E-05	N	Uncertain	na
<i>COL2A1</i> (44)	chr12:48,371,857	C>T	NM_001844.5: c.3047G>A rs146046296	NP_001835.3: p.Arg1016Lys	missense	VUS (PM2,PM1, PP2,BP6)	1.22E-04	Y	Uncertain	conflict
<i>MMP25</i> (3)	chr16:3,100,129	C>T	NM_022468.5: c.352C>T rs149856253	NP_071913.1: p.Arg118*	nonsense	VUS (PM2)	5.70E-04	Y	na	na
<i>RARA</i> [#] (7)	chr17:38,510,757	A>C	NM_000964.4: c.1011A>C	NP_000955.1: p.Gly337Gly	splice-site	VUS (PM2,PP3)	na	Y	Deleterious	na
<i>MMP9</i> (6)	chr20:44,640,347	T>A	NM_004994.3: c.958T>A rs202151407	NP_004985.2: p.Tyr320Asn	missense	VUS (PM2)	3.30E-05	Y	Uncertain	na
<i>TBX22</i> (4)	chrX:79,279,653	A>C	NM_001109878.2: c.448A>C rs1156605251	NP_001103348.1: p.Lys150Gln	missense	VUS (PM2)	9.23E-05	Y	Uncertain	na

[#], Variant detected in proband with Pierre-Robin sequence; Y/N, yes (Y) or no (N); na, not available; conflict, conflicting interpretations of pathogenicity (uncertain significance, likely benign or benign); Ref>Alt, reference and alternative allele, where reference allele is listed first

^a, All listed variants are heterozygous and were all found in one family, with the exception of rs117056541 (*TBX10*), which was identified in two families; ^b, The frequency of the alternative allele from the Genome Aggregation Database (GnomAD); ^c, Predicted deleterious by *in silico* tools – missense SNV predicted deleterious by at least three *in silico* prediction tools (SIFT, PolyPhen-2, MutationAssessor, MutationTaster, FATHMM, CADD, MetaLR, REVEL or GERP++), splice site SNV predicted to alter splicing by at least one splice site prediction tool (MaxEntScan, dbSCSNV Ada, or SpliceAI), and nonsense SNV predicted deleterious by two tools (MutationTaster and CADD); ^d, Interpretation in ClinVar database

Table S4. Primer sequences and parameters used for Sanger sequencing of protein-coding exons and flanking intronic regions of *GRHL3*, *IRF6*, and *TBX22*

		Primer pairs' sequences	Genomic location (GRCh38/hg38)	Size [bp]	Ta [°C]	c Mg ²⁺ [mM]	c F/R [μM]
<i>GRHL3</i>*							
Exon 1	F	5'-AGTTATCTGTTTGTCCCTGTACT-3'	Chr1:24,319,155-	758	58.9	2.5	0.4
	R	5'-AGAAAGCCTCGATGGGATTG-3'	24,319,912				
Exon 1'	F	5'-ACGGAAGCACTGGGATCTTA-3'	Chr1:24,322,951-	237	56.4	2.5	0.4
	R	5'-TTCCCCTTAACGACTCCAAA-3'	24,323,187				
Exon 2	F	5'-TTGACTCTCCTTACTTGCATTTCAG-3'	Chr1:24,331,402-	232	62.4	3.5	1.0
	R	5'-GTCCAGAGGGGGTAGACGTA-3'	24,331,633				
Exon 3	F	5'-GATTGAGGCTCCTACCAGCA-3'	Chr1:24,334,583-	225	62.4	2.0	0.2
	R	5'-CTGTGCCTCCTGTGCCTAAT-3'	24,334,807				
Exon 4	F	5'-GATCAAAGGCCAGTGTGTCA-3'	Chr1:24,336,385-	579	66.9	2.0	0.4
	R	5'-GTACATGGCAATCCCGTCTC-3'	24,336,963				
Exon 5	F	5'-GAGGCTTCCCAGAGTGAATG-3'	Chr1:24,337,021-	228	62.4	2.0	0.4
	R	5'-TGATTGTGCCTGCTATGCTC-3'	24,337,248				
Exon 6	F	5'-TGGTTTTGTACCCCTCAAG-3'	Chr1:24,337,388-	573	62.4	2.0	0.2
	R	5'-GCCTCTTCCTAGAGCCTGGT-3'	24,337,960				
Exon 7	F	5'-CATTGGAACCAGGCTCTAGG-3'	Chr1:24,337,934-	199	62.4	2.0	0.4
	R	5'-CTGGAATGGAGGCTTGAGAA-3'	24,338,132				
Exon 8	F	5'-CTCCTTCTGGTCTCCTGTGG-3'	Chr1:24,339,642-	225	62.4	2.0	0.4
	R	5'-CGTCGGAGCAAATGACACTA-3'	24,339,866				
Exon 9	F	5'-TGTACCAACAGCGGTGAAAA-3'	Chr1:24,341,837-	498	62.4	2.0	0.2
	R	5'-GTCAGGGTTTGGCCCTTCTA-3'	24,342,334				
Exon 10	F	5'-CTTCTGTCTGTCCGCCTCTC-3'	Chr1:24,342,464-	485	62.4	2.0	0.2
	R	5'-GCCGAAGGTAGGTCGTCTC-3'	24,342,948				
Exon 11	F	5'-GTGAAGGGGAGAAGGAGACC-3'	Chr1:24,342,811-	545	66.9	2.0	0.2
	R	5'-ACAGTGAGCACCACAAGTGC-3'	24,343,355				
Exon 12	F	5'-TCCACTCAACATCTCAGGTCA-3'	Chr1:24,344,716-	246	62.4	1.0	0.4
	R	5'-AGGGAGGCGTGTGGTTGT-3'	24,344,961				
Exon 13	F	5'-GCATCCACCCTTGTGTTTTTC-3'	Chr1:24,346,437-	243	62.4	2.0	0.4
	R	5'-GTGGGCCTGGTGAGTAAATG-3'	24,346,679				
Exon 14	F	5'-CCCCTGAGATGATCCTGTTC-3'	Chr1:24,347,407-	240	62.4	2.5	0.4
	R	5'-CCTCCTGTGGGAAATCAGAA-3'	24,347,646				
Exon 15	F	5'-CAGCAGGCAATGAATCACC-3'	Chr1:24,350,023-	247	62.4	1.0	0.4
	R	5'-GCAGATGGAGAGGTGACCAG-3'	24,350,269				
Exon 16	F	5'-GTGCCTAAAATGTGGCAGGT-3'	Chr1:24,354,257-	682	62.4	2.0	0.2
	R	5'-CTGACCAAAGGTTGAGAAATCC-3'	24,354,938				
Exon 16'	F	5'-ACTTTCCTGCAAACCTCGAA-3'	Chr1:24,364,134-	396	65.1	2.0	0.4
	R	5'-CCTTCCCTTAAAGAGCTAGGC-3'	24,364,529				
<i>IRF6</i>							
Exon 3	F	5'-TCTAGTAGATGGGAAAGGTGGC-3'	Chr1:209,801,133	373	62.4	2.0	0.4
	R	5'-AGCATTCTCTCTGTTTCACCAG-3'	- 209,801,505				
Exon 4	F	5'-TCTGTAAATCGGGGTGGGA-3'	Chr1:209,796,223	498	62.4	2.0	0.4
	R	5'-AGGCTTTCTTGCTTTATCCATCT-3'	- 209,796,720				
Exon 5	F	5'-TGTACTGAACCTGAGGAGCC-3'	Chr1:209,795,136	399	62.4	2.0	1.0
	R	5'-TCACCTCTGACTCCCACTTG-3'	- 209,795,534				
Exon 6	F	5'-CTGATGGGCTGGAAGATCAC-3'	Chr1:209,792,125	440	58.9	2.0	0.2
	R	5'-GCTAGCCAGGAAACAGAAACA-3'	- 209,792,564				
Exon 7	F	5'-CTGGTTGAAAGGTGGCTTGA-3'	Chr1:209,790,393-	597	62.4	2.0	0.4
	R	5'-TTTGCCATGCCAGGAAAGC-3'	209,790,989				
Exon 8	F	5'-TCAATGATGTCCAAGAGAACCC-3'	Chr1:209,789,571-	471	65.1	2.0	0.2
	R	5'-GGGCTGATGGATGCTTGATG-3'	209,790,041				
Exon 9	F	5'-ACAGTGAGCCTTGGCAGAA-3'	Chr1:209,788,234-	587	65.1	2.0	1.0
	R	5'-AGCAAAGTCTGAAGGGTGATTT-3'	209,788,820				
<i>TBX22</i>							
Exon 2	F	5'-CTCCCTCCCTAACCCAGTTC-3'	ChrX:80,022,206-	496	65.1	2.5	0.2
	R	5'-AGAGCTTTCGCGAAGGTAG-3'	80,022,701				

Exon 3	F	5'-AGTGGGCATGTGAACTGTGA-3'	ChrX:80,023,001-	395	58.9	2.5	0.2
	R	5'-TGAGGGGCTATGGCTTTCTA-3'	80,023,395				
Exon 4	F	5'-GGGTATTGTGAGTCCCTTCA-3'	ChrX:80,023,951-	250	58.9	2.5	0.4
	R	5'-CCTGGGCATAGAGTACCATTC-3'	80,024,200				
Exon 5	F	5'-CTGCACCTAATGCCACAGC-3'	ChrX:80,025,570-	238	66.9	2.0	0.8
	R	5'-TCCTCGGGGTAAAGGATTGT-3'	80,025,807				
Exons 6 and 7	F	5'-GGAAGTAGGGTTTGGGACTGA-3'	ChrX:80,026,644-	700	65.1	2.0	0.4
	R	5'-GATGTCATTGCTATGCTGCTT-3'	80,027,343				
Exon 8	F	5'-TGGGGATGCTGAAAGTTGA-3'	ChrX:80,027,959-	247	54.4	2.0	0.2
	R	5'-CGCAGTTACCCATAAGTACACATT-3'	80,028,205				
Exon 9	F	5'-GCAGGAACATCAAATGTCAAG-3'	ChrX:80,030,444-	651	58.9	2.5	0.4
	R	5'-TTGCTGGATACCAATGAACA-3'	80,031,094				

*, *GRHL3* exons listed based on the canonical transcript (alternative exons found in other transcripts are marked with »'«)

Size, amplicon size; Ta, optimal annealing temperature; c Mg²⁺, optimal Mg²⁺ concentration; c F/R, optimal primer concentration

Table S5. 72 additional genes used for the second step of the sequence analysis

Gene symbol	Gene name
<i>ABCA4</i>	ATP Binding Cassette Subfamily A Member 4
<i>ADH1C</i>	Alcohol Dehydrogenase 1C (Class I), Gamma Polypeptide
<i>APC</i>	APC Regulator Of WNT Signaling Pathway
<i>ARHGAP29</i>	Rho GTPase Activating Protein 29
<i>ASS1</i>	Argininosuccinate Synthase 1
<i>AXIN2</i>	Axin 2
<i>BCL3</i>	BCL3 Transcription Coactivator
<i>BHMT</i>	Betaine--Homocysteine S-Methyltransferase
<i>BMP2</i>	Bone Morphogenetic Protein 2
<i>BMP4</i>	Bone Morphogenetic Protein 4
<i>BRIP1</i>	BRCA1 Interacting Helicase 1
<i>BSPRY</i>	B-Box And SPRY Domain Containing
<i>CCL2</i>	C-C Motif Chemokine Ligand 2
<i>CDH1</i>	Cadherin 1
<i>CLPTM1</i>	CLPTM1 Regulator Of GABA Type A Receptor Forward Trafficking
<i>COL2A1</i>	Collagen Type II Alpha 1 Chain
<i>COL11A2</i>	Collagen Type XI Alpha 2 Chain
<i>CTNNB1</i>	Catenin Beta 1
<i>DHFR</i>	Dihydrofolate Reductase
<i>DLG1</i>	Discs Large MAGUK Scaffold Protein 1
<i>DVL2</i>	Dishevelled Segment Polarity Protein 2
<i>EDN1</i>	Endothelin 1
<i>FGF1</i>	Fibroblast Growth Factor 1
<i>FGF2</i>	Fibroblast Growth Factor 2
<i>FGF3</i>	Fibroblast Growth Factor 3
<i>FGF10</i>	Fibroblast Growth Factor 10
<i>FGFR1</i>	Fibroblast Growth Factor Receptor 1
<i>FN1</i>	Fibronectin 1
<i>FOXE1</i>	Forkhead Box E1
<i>GCH1</i>	GTP Cyclohydrolase 1
<i>GNMT</i>	Glycine N-Methyltransferase
<i>GREM1</i>	Gremlin 1, DAN Family BMP Antagonist
<i>GSK3B</i>	Glycogen Synthase Kinase 3 Beta
<i>JAG2</i>	Jagged Canonical Notch Ligand 2
<i>MAFB</i>	MAF BZIP Transcription Factor B
<i>MMP3</i>	Matrix Metalloproteinase 3
<i>MMP9</i>	Matrix Metalloproteinase 9
<i>MMP13</i>	Matrix Metalloproteinase 13
<i>MMP25</i>	Matrix Metalloproteinase 25
<i>MSX1</i>	Msh Homeobox 1
<i>MTHFD1</i>	Methylenetetrahydrofolate Dehydrogenase, Cyclohydrolase And Formyltetrahydrofolate Synthetase 1
<i>MTHFR</i>	Methylenetetrahydrofolate Reductase
<i>MTR</i>	5-Methyltetrahydrofolate-Homocysteine Methyltransferase
<i>MTRR</i>	5-Methyltetrahydrofolate-Homocysteine Methyltransferase Reductase
<i>NECTIN1</i>	Nectin Cell Adhesion Molecule 1
<i>NECTIN2</i>	Nectin Cell Adhesion Molecule 2
<i>NOG</i>	Noggin

<i>PAH</i>	Phenylalanine Hydroxylase
<i>PCYT1A</i>	Phosphate Cytidyltransferase 1A, Choline
<i>PIAS1</i>	Protein Inhibitor Of Activated STAT 1
<i>PIAS2</i>	Protein Inhibitor Of Activated STAT 2
<i>PON1</i>	Paraoxonase 1
<i>PTCH1</i>	Patched 1
<i>PVR</i>	PVR Cell Adhesion Molecule
<i>RARA</i>	Retinoic Acid Receptor Alpha
<i>RING1</i>	Ring Finger Protein 1
<i>SKI</i>	SKI Proto-Oncogene
<i>SUMO1</i>	Small Ubiquitin Like Modifier 1
<i>SUMO3</i>	Small Ubiquitin Like Modifier 3
<i>TBX10</i>	T-Box Transcription Factor 10
<i>TCN2</i>	Transcobalamin 2
<i>TGFA</i>	Transforming Growth Factor Alpha
<i>TGFB1</i>	Transforming Growth Factor Beta 1
<i>TGFB3</i>	Transforming Growth Factor Beta 3
<i>TIMP2</i>	TIMP Metalloproteinase Inhibitor 2
<i>TIMP3</i>	TIMP Metalloproteinase Inhibitor 3
<i>VAX1</i>	Ventral Anterior Homeobox 1
<i>WNT3</i>	Wnt Family Member 3
<i>WNT3A</i>	Wnt Family Member 3A
<i>WNT5A</i>	Wnt Family Member 5A
<i>WNT8A</i>	Wnt Family Member 8A
<i>WNT9B</i>	Wnt Family Member 9B

Table S6. Primer sequences and parameters used for variant validation and co-segregation analysis by Sanger sequencing

Gene [exon] (variant)*	Primer pairs' sequences	Genomic location (GRCh38/hg38)	Size [bp]	Ta [°C]	c Mg ²⁺ [mM]	c F/R [μM]
<i>MTHFR</i> [7] (rs200138092)	F 5'-CGCTCCCAAGAACAAGATG-3' R 5'-CCCTCTAGCCAATCCCTTGT-3'	Chr1:11,794,679- 11,794,912	234	56.4	2.0	0.2
<i>MTHFR</i> [6] (rs771684838)	F 5'-ATGGCACAGACATCCATCTG-3' R 5'-CCCAGGTGTGCTCCTAGAAT-3'	Chr1:11,794,999- 11,795,495	497	62.4	2.0	0.2
<i>GRHL3</i> [10] (c.1285G>T)	F 5'-CTTCTGTCTGTCCGCCTCTC-3' R 5'-GCCGAAGGTAGGTCGTCTC-3'	Chr1:24,342,464- 24,342,948	485	62.4	2.0	0.2
<i>IRF6</i> [3] (rs121434229)	F 5'-TCTAGTAGATGGGAAAGGTGGC-3' R 5'-AGCATTCTCTCTGTTTCACCAG-3'	Chr1:209,801,133- 209,801,505	373	62.4	2.0	0.4
<i>IRF6</i> [6] (c.622C>T)	F 5'-CTGATGGGCTGGAAGATCAC-3' R 5'-GCTAGCCAGGAAACAGAAACA-3'	Chr1:209,792,125- 209,792,564	440	58.9	2.0	0.2
<i>IRF6</i> [7] (c.687delG)	F 5'-CTGGTTGAAAGGTGGCTTGA-3' R 5'-TTTGCCATGCCAGGAAAGC-3'	Chr1:209,790,393- 209,790,989	597	62.4	2.0	0.4
<i>IRF6</i> [9] (rs1553247595)	F 5'-ACAGTGAGCCTTGGCAGAA-3' R 5'-AGCAAAGTCTGAAGGGTGATTT-3'	Chr1:209,788,234- 209,788,820	587	65.1	2.0	1.0
<i>DLG1</i> [20] (rs758466083)	F 5'-GGCTTCACAAACCTTCTTGT-3' R 5'-GGAGTAGCACATAGCATCACAAA-3'	Chr3:197,066,692- 197,067,041	350	62.4	2.0	0.2
<i>FGFR1</i> [4] (rs138489552)	F 5'-GGAGGTGTCTTGCCCATCT-3' R 5'-GAGGGTGTCCGTGTTCATCT-3'	Chr8:38,428,177- 38,428,527	351	62.4	2.0	0.2
<i>TBX10</i> [3] (rs117056541)	F 5'-CCCGCACCTGTATCTCTTGT-3' R 5'-GAAGCCTCTGTGGGAGGAAT-3'	Chr11:67,634,809- 67,635,058	250	62.4	2.0	0.2
<i>MMP3</i> [4] (rs782241657)	F 5'-CCCCTAGATCACAAAGATCCA-3' R 5'-ACTGGAAAAGGCTCAAAAGG-3'	Chr11:102,841,900 -102,842,342	443	62.4	2.0	0.2
<i>COL2A1</i> [44] (rs146046296)	F 5'-GAAGCAGCCCTTGGTCTCTA-3' R 5'-GTCGGTGAGTGTGGGGTATC-3'	Chr12:47,977,813- 47,978,294	482	66.9	2.0	0.2
<i>BMP4</i> [3] (rs121912767)	F 5'-ACTGGGGGAAGAGACTGACC-3' R 5'-GACTTCGAGGCGACACTTCT-3'	Chr14:53,951,834- 53,952,057	224	62.4	1.0	0.4
<i>JAG2</i> [24] (c.3004A>G)	F 5'-CCAGGGATGTAGGTGTCAGG-3' R 5'-GAGCCTCTGTCCAGCATAGC-3'	Chr14:105,144,848 -105,145,124	277	65.1	2.0	0.2
<i>MMP25</i> [3] (rs149856253)	F 5'-GCAGGCTCTCCTATTGTGGA-3' R 5'-GGCAAAGTCGATGAGGATGT-3'	Chr16:3,049,958- 3,050,410	453	62.4	2.0	0.2
<i>RARA</i> [7] (c.1011A>C)	F 5'-AGAGTGGGTTCGGGTTCAGT-3' R 5'-CCCTTGCTGTGTAGTTGCAG-3'	Chr17:40,354,253- 40,354,644	392	66.9	2.0	0.2
<i>MMP9</i> [6] (rs202151407)	F 5'-ATGAGAGATGGGATGAACTGC-3' R 5'-CAAAGTATGGGAGGGAAGA-3'	Chr20:46,011,460- 46,011,834	375	65.1	2.0	0.2
<i>TBX22</i> [4] (rs1156605251)	F 5'-GGGTATTGTGAGTCCCTTCA-3' R 5'-CCTGGGCATAGAGTACCATTTC-3'	ChrX:80,023,951- 80,024,200	250	58.9	2.5	0.4
<i>TBX22</i> [9] (rs761750322)	F 5'-GCAGGAACATCAAATGTCAAG-3' R 5'-TTGCTGGATACCAATGAACA-3'	ChrX:80,030,444- 80,031,094	651	58.9	2.5	0.4

*, Variants without 'rsID' number were named by the position of the nucleotide change, determined based on the Ensembl canonical transcripts

Size, amplicon size; Ta, optimal annealing temperature; c Mg²⁺, optimal Mg²⁺ concentration; c F/R, optimal primer concentration

Table S7. Primer sequences and parameters used for real-time quantitative PCR analysis of *TBX22* coding exons

	Primer pairs' sequences	Genomic location (GRCh38/hg38)	Amplicon size (bp)	Ta (°C)	c F/R (nM)
<i>TBX22</i>					
Exon 2	F 5'-CTCCCTCCCTAACCCAGTTC-3' R 5'-AGAGCTTTCCGCGAAGGTAG-3'	ChrX:80,022,206- 80,022,701	496	64	100
Exon 3	F 5'-CGGAGCTTGGGGAATCTG-3' R 5'-GGTAAGTCCGAGCCCTATCA-3'	ChrX:80,022,922- 80,023,417	496	64	400
Exon 4	F 5'-GGGTATTGTGAGTCCCTTCA-3' R 5'-CCTGGGCATAGAGTACCATTTC-3'	ChrX:80,023,951- 80,024,200	250	59	400
Exon 5	F 5'-CTGCACCTAATGCCACAGC-3' R 5'-TCCTCGGGGTAAAGGATTGT-3'	ChrX:80,025,570- 80,025,807	238	64	100
Exon 6	F 5'-TGGGACTGAAGCCAGTTTTT-3' R 5'-TTTCCATTTCGCCTGAGACAT-3'	ChrX:80,026,656- 80,026,900	245	60	200
Exon 7	F 5'-AACGTGTTTCAACATCTCTTCTGG-3' R 5'-TATTAGATGTCATTGCTATGCTGC-3'	ChrX:80,027,149- 80,027,348	200	60	200
Exon 8	F 5'-GAAATTGCATTCTGGGGATG-3' R 5'-AAGGGTTCCGGATGTTGTTA-3'	ChrX:80,027,947- 80,028,142	196	60	200
Exon 9	F 5'-GCAGGAACATCAAATGTCAAG-3' R 5'-TTGCTGGATACCAATGAACA-3'	ChrX:80,030,444- 80,031,094	651	59	400
<i>G6PD</i>					
Exon 3	F 5'-TCGGGGCTCTTCTGTCTGTA-3' R 5'-CGATGAAGGTGTTTTCGGGC-3'	ChrX:154,536,002- 154,536,332	331	64	400
<i>IRF6</i>					
Exon 5	F 5'-TGTACTGAACCTGAGGAGCC-3' R 5'-TCACCTCTGACTCCCACTTG-3'	Chr1:209,795,136- 209,795,534	399	64	100

Ta, optimal annealing temperature; c F/R, optimal primer concentration

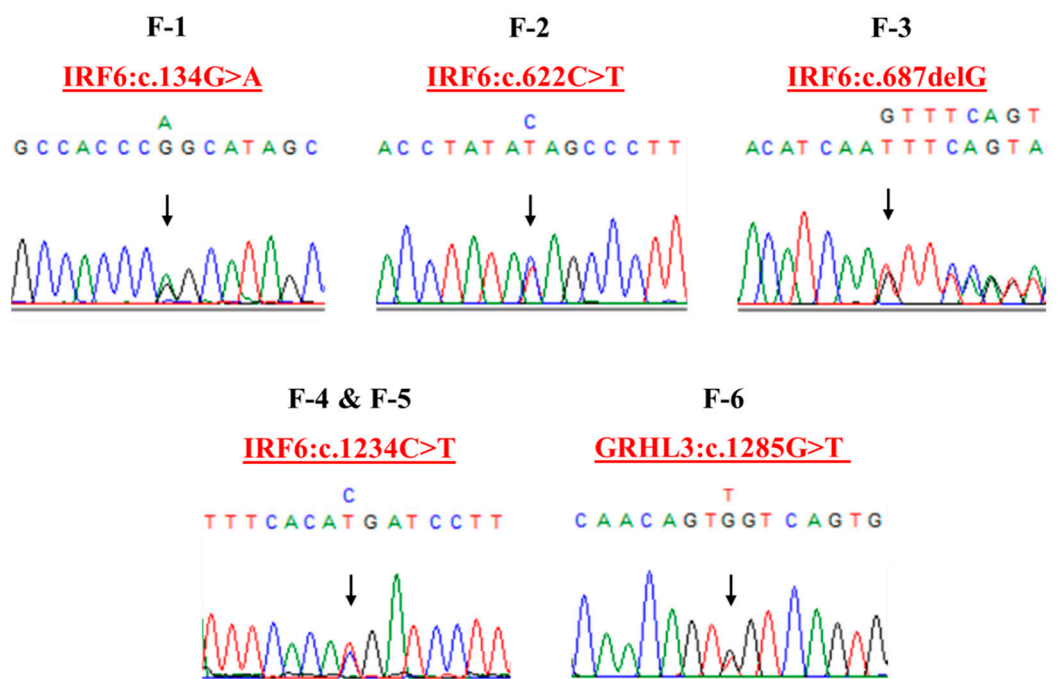


Figure S1. Sequencing chromatograms of *IRF6* and *GRHL3* variants confirming the diagnosis of Van der Woude syndrome in six families (F-1 to F-6). Each base affected by a variant is marked with an arrowhead.

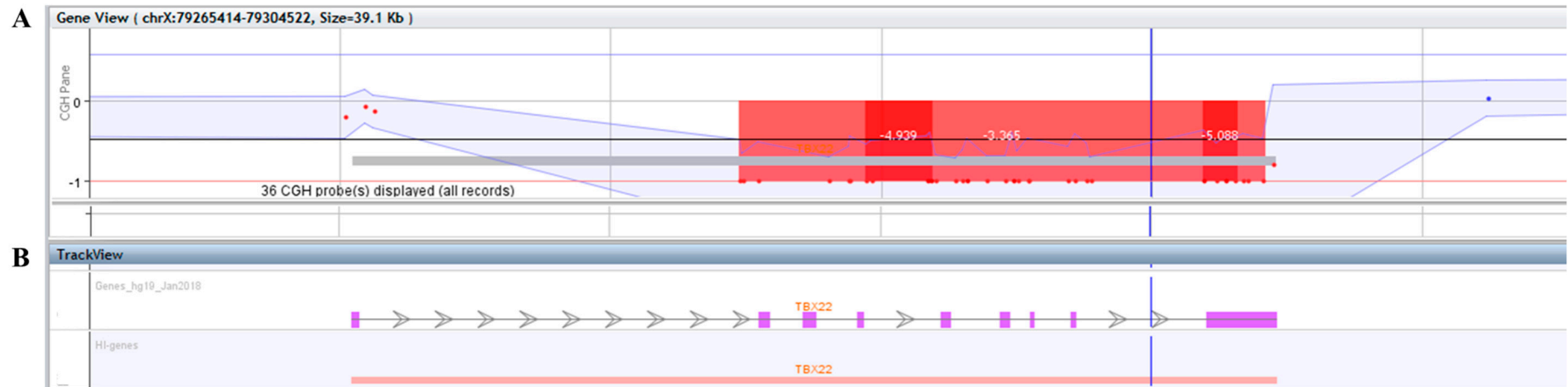


Figure S2. The result of microarray-based comparative genomic hybridization (array CGH) analysis in the proband of family 7 (F-7) with X-linked cleft palate with or without ankyloglossia. It shows the deletion (arr[GRCh37] Xq21.1(79,277,377_79,287,288)x0) encompassing exons 2-9 of the *TBX22* gene. (A), The log₂ ratio values of the probes covering the region of the *TBX22* gene, with the red area depicting the deleted region. To the left and right of the deleted region are the adjacent signals with normal copy number values; (B), The canonical transcript of the gene in the deleted region, the *TBX22* gene, with the purple squares representing exons and the flanking grey lines representing introns; the pink line shows that the *TBX22* gene is haploinsufficient.