Journal of Human Genetics (2015) 60, 17–25

www.nature.com/jhg

ORIGINAL ARTICLE

Investigation of genetic factors underlying typical orofacial clefts: mutational screening and copy number variation

Milena Simioni[1], Tânia Kawasaki Araujo[1], Isabella Lopes Monlleo[2], Cláudia Vianna Maurer-Morelli[1] and Vera Lúcia Gil-da-Silva-Lopes[1]

Typical orofacial clefts (OFCs) comprise cleft lip, cleft palate and cleft lip and palate. The complex etiology has been postulated to involve chromosome rearrangements, gene mutations and environmental factors. A group of genes including *IRF6*, *FOXE1*, *GLI2*, *MSX2*, *SKI*, *SATB2*, *MSX1* and *FGF* has been implicated in the etiology of OFCs. Recently, the role of the copy number variations (CNVs) has been studied in genetic defects and diseases. CNVs act by modifying gene expression, disrupting gene sequence or altering gene dosage. The aims of this study were to screen the above-mentioned genes and to investigate CNVs in patients with OFCs. The sample was composed of 23 unrelated individuals who were grouped according to phenotype (associated with other anomalies or isolated) and familial recurrence. New sequence variants in *GLI2*, *MSX1* and *FGF8* were detected in patients, but not in their parents, as well as in 200 control chromosomes, indicating that these were rare variants. CNV screening identified new genes that can influence OFC pathogenesis, particularly highlighting *TCEB3* and *KIF7*, that could be further analyzed. The findings of the present study suggest that the mechanism underlying CNV associated with sequence variants may play a role in the etiology of OFC.
Journal of Human Genetics (2015) **60,** 17–25; doi:10.1038/jhg.2014.96; published online 13 November 2014

INTRODUCTION

Orofacial clefts (OFCs) comprise any cleft (that is, a break or gap) in the orofacial structures. Most of the research on OFCs has primarily focused on typical cases such as cleft lip (CL), cleft palate (CP) and cleft lip and palate (CLP).[1] The prevalence of CLP is 6.64 per 10 000, CL is 3.28 per 10 000,[2] and CP is 4.50 per 10 000 live births.[3] Most cases of OFCs occur as an isolated defect, although they can be associated with other anomalies or as part of syndromes.[4]

The complex etiology of OFCs involves chromosome rearrangements, gene mutations and environmental factors.[5,6] It has been suggested that OFCs are caused by genetic variations in more than one gene because several processes are involved in lip and palate formation including cell proliferation, differentiation, adhesion and apoptosis.[7]

A group of genes including *IRF6*, *FOXE1*, *GLI2*, *MSX2*, *SKI*, *SATB2*, *MSX1* and *FGF* has been identified to contribute to OFC etiology. Mutations in *IRF6* have been detected in 12% of OFC cases,[8] those in *FOXE1*, *GLI2*, *MSX2*, *SKI*, *SATB2* and *SPRY2* account for 6%,[9] those in *MSX1* are responsible for 2%[10,11] and those in the *FGF* family of genes, mainly *FRGR1* and *FGF8*, contribute to 3% of the cases.[12] These genes represent only a small proportion of known genetic factors involved in the development of OFCs.[7] Despite efforts to understand OFC etiology, the molecular mechanisms underlying cleft development have not been fully characterized.

Studies using array-based techniques have uncovered various large-scale copy number variations (CNVs), deletions and duplications that substantially contribute to human genomic variation.[13,14] In addition, CNVs play a role in genetic defects and diseases by modifying gene expression, disrupting gene sequence or altering gene dosage.[15–17] CNV screening has proven to be a powerful strategy in identifying candidate genes and/or chromosome regions involved in various disorders including OFC.[18–20]

To investigate the genetic aspects of OFC, the aims of this study were to screen *IRF6*, *FOXE1*, *GLI2*, *MSX2*, *SKI*, *SATB2*, *MSX1*, *FGF8* and *FGFR1* by Sanger sequencing and to investigate the role of CNVs by array genomic hybridization (aGH) in a clinically well-characterized group of patients with OFC.

MATERIALS AND METHODS

All patients or parents representing their child provided their informed consent, as required and approved by the Research Ethics Committee of our institution (#714/2008).

Patients were ascertained at the Clinical Hospital, University of Campinas, and at the Faculty of Medicine, Federal University of Alagoas. The study

[1]Department of Medical Genetics, Faculty of Medical Sciences, University of Campinas, UNICAMP, Campinas, São Paulo, Brazil and [2]Clinical Genetics Service, Faculty of Medicine, University Hospital and Genetics Sector, Federal University of Alagoas, Maceió, Alagoas, Brazil
Correspondence: Professor VL Gil-da-Silva-Lopes, Department of Medical Genetics, University of Campinas, Tessália Vieira de Camargo Street, 126-CEP, Campinas, São Paulo 13083-887, Brazil.
E-mail: vlopes@fcm.unicamp.br
Received 1 September 2014; revised 1 October 2014; accepted 10 October 2014; published online 13 November 2014

population was composed of 23 unrelated individuals who were grouped according to phenotype (associated with other anomalies or isolated) and familial recurrence (Table 1). OFCs associated with congenital defects included seven patients with CLPs, five with CLs and three with CPs. In the isolated OFC group, three patients showed CLP, four CL and one CP. Four familial cases of CLP were identified: patient 7 has an aunt with isolated CL, patient 16 has an uncle with isolated CLP, the mother of patient 18 presented with CLP and the mother of patient 23 was affected by isolated CP.

All patients were evaluated by clinical geneticists using the same clinical protocol and classified according to *International Perinatal Database of Typical Oral Clefts* (2011), as well as by echocardiography. In addition, GTG-banding karyotypes (600 bands) of the patients were prepared. The karyotypes were normal in 21 cases, patient 11 was 46,XX,t(4;5)(p10;p10)pat(20) and patient 14 had a chromosomal constitution of 46,XY,ins(11;?)(p13;?)(20).

Mutation screening for the entire coding region and exon–intron boundaries was performed for the following candidate genes: *IRF6*, *FOXE1*, *GLI2*, *MSX2*, *SKI*, *SATB2*, *MSX1* and *FGF8*. For *FGFR1*, only three mutations previously described by Riley *et al.*[12] were analyzed. Sequencing was performed in an Applied Biosystems 3500xL Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) using a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). To predict the effects of the identified variants, computational algorithms such as Grantham score, Panther, PolyPhen, SIFT and SNP&GO were applied. In addition, a group of 100 Brazilian control individuals without OFC and encompassing three generations were sequenced for variants detected.

The CNV pattern of the patients was determined by the aGH technique using the Affymetrix Genome-Wide Human single-nucleotide polymorphism (SNP) Array 6.0 (Affymetrix, Santa Clara, CA, USA) according to the manufacturer's instructions. CNV analysis of trios (patient–parents) was feasible to perform in patients 2, 4, 8, 9, 12, 14, 17, 19, 21, 22 and 23. In contrast, CNV analysis of patients and mothers was possible for individuals 3, 5, 6, 10, 11 and 20. In familial cases, the patient and the affected relative were analyzed.

Data analysis was performed using the Genotyping Console v. 3.0.2 (HMM) (Affymetrix) software. Comparisons were conducted using three different strategies as follows: patients vs 20 Brazilian control individuals without OFCs in three generations; patients vs 50 Brazilian control individuals; and patients vs the HapMap control data set. For CNV screening, regions of sizes $\geqslant$ 300 kb and involving 25 markers for deletion and 50 markers for duplications were used. CNVs of sizes $<$ 300 kb were carefully verified to detect genes related to OFCs or new ones that could be related.

RESULTS

Sequence screening of candidate genes identified several SNPs that have been previously reported in a public database (http://www.ncbi.nlm.nih.gov/SNP), as well as three undocumented sequence variants. Patient 21 presented one variant in *GLI2* (c.2341C$>$T, p.Leu761Phe) (Figure 1a). Patient 8 harbored a variant in *MSX1* (c.329C$>$T, Ala32Val) (Figure 1b). Patient 6 showed an *FGF8* variant defined as c.765C$>$A, p.Glu236Lys (Figure 1c). None of these variants were detected in the control group of 100 Brazilians. The results of five protein prediction programs varied in terms of the effects of these alterations. Panther and Polyphen considered these variants tolerant and benign. However, SFIT considered them intolerant (score 0.00). Gratham and SNP&GO pointed *GLI2* variant as conservative (escore 22) and deleterious (reability = 10); *MSX1* and *FGF8* moderately conservative (escore 64 and 56, respectively) and neutral (reability = 1 and reability = 9, respectively).

According to previously defined parameters of size and markers, the role of CNVs detected by the three analytical approaches is different and summarized in Supplementary Table 1. CNVs that were detected in all three analytical tests were selected to gene identification, research in Database of Chromosomal Imbalance and Phenotype in Humans (DECIPHER) and Database of Genomic Variants (DGV) (Table 2). DECIPHER cases were considered relevant up to four CNVs beyond the one overlapping with CNV presented by our patient or if OFC was part of the phenotype.

CNVs of $<$300 kb in size were also assessed; however, only two were considered relevant based on genes involved. In patient 3, a 96-kb duplication was detected in *FGFR1* that is located in the chromosomal region 8p12 (nt 38 431 900–38 441 500 (hg18)). Patient 8 showed a 270-kb deletion in the chromosomal region 1p36.11 (nt 23 903 625–24 173 440 (hg18)) that encompassed 8 genes including *TCEB3* (Figure 2). Both CNVs were confirmed using the three types of analysis.

Patient 14 presented a karyotype of 46,XY,ins(11;?)(p13;?)[20]. The aGH analysis detected two duplications: a 17.09-Mb segment at the chromosomal region 15q25–q26 (nt 81 869 248–98 962 477 bp (hg18)) and a 3.8-Mb duplication at the chromosomal region 8p23.1 (nt 8 129 435–11 934 586 bp (hg18)). The complete report of this patient has been published elsewhere,[21] although the present study highlights the 15q–15q26 region that involves *KIF7* (Figure 3). Patient 11, who showed a karyotype of 46,XX,t(4;5)(p10;p10), had no alterations involving breakpoint regions that would further characterize it as a balanced translocation.

DISCUSSION

Our strategy to investigate the genetic factors involved in OFCs was based on standard clinical evaluation, mutational screening and aGH analysis. The main idea of this research was to perform a screening of variants (sequence and copy number) involved in orofacial clefts that justifies the sample composed of different types of clefts (CLP, CP and CL).

Undocumented variants were detected in patients with *GLI2*, *MSX1* and *FGF8*, whereas these were not observed in their parents as well as in the 200 control chromosomes, indicating that these were rare variants. The predicted effects at the protein level using *in silico* algorithms showed discordant results. Functional studies are therefore necessary to elucidate how these variants affect gene expression and protein production during development and to establish their role in OFC etiology.

GLI2 belongs to a zinc-finger protein class that is required for the expression of genes during embryogenesis and is involved in the sonic hedgehog signaling process.[22,23] Mutations in *GLI2* together with *FOXE1*, *MSX2*, *SKI* and *SATB2* have been detected in 6% of OFC cases.[9] Mutations in this gene have also been reported in patients with OFC, holoprosencephaly and facial anomalies.[24,25] The *GLI2* variant was detected in patient 21, who was classified as an isolated CL case.

MSX1 encodes a member of the muscle segment homeobox gene family that controls gene expression during the development of palatal shelves.[26] This gene is also involved in epithelial–mesenchymal growth and differentiation of specific tissues. Animal models of growth disruption because of mutations in *MSX1* have been shown to develop palatal clefts.[27] Mutations in *MSX1* contribute to 2% of isolated OFC cases that consist of patients of different ethnicities.[10] Patient 8, who carried this variant, presented a CL that was associated with other anomalies as well as CNV (to be discussed later) that might have played a role in OFC development.

Mammalian fibroblast growth factors (*FGF1–FGF10* and *FGF16–FGF23*) control a wide spectrum of biological functions during development and adult life.[28] *FGF8* expression occurs during gastrulation as well as during the development of the brain, heart, limbs and craniofacial structures including labial and palatal shelves.[29–32] Sequence screening of 12 fibroblast growth factor genes (*FGFR1*, *FGFR2*, *FGFR3*, *FGF2*, *FGF3*, *FGF4*, *FGF7*, *FGF8*, *FGF9*, *FGF10*,

Table 1 Clinical aspects of patients

ID	*Sex*	*Cleft type*	*Craniofacial aspects*	*Other malformations*
1	F	ACLP	Flat occipital, low frontal hairline, malar hypoplasia, low-set and malformed ears, synophrism, up-slanting palpebral fissures, asymmetrical nostrils, long and thin *philtrum*, bifid uvula	Webbed neck, *pectus excavatum*, sacral dimple, *spina bifida*, clinodactyly and syndactyly of hands, clinodactyly and brachydactyly of feet, diffused hypochromic spots, generalized muscle atrophy
2	M	ACLP	Epicanthus	Wide-spaced and hypoplastic nipples, a single bilateral palmar crease
3	M	ACLP	Epicanthus, oblique palpebral fissures down	Clinodactyly of hands, splenomegaly, cardiac defects, delayed psychomotor development
4	F	ACLP	Prominent frontal, high frontal hairline, long face, malar hypoplasia, ocular hypertelorism, blepharophimosis, atresia of lacrimal ducts	Thin skin, sparse, curly and fine hair, conical teeth, progressive hearing loss
5	F	ACLP	Low-set and malformed ears, long face, palpebral ptosis, microcephaly, prognathism	Congenital cardiac defects, delayed psychomotor development, behavioral disturbance, short stature, sensorineural deafness, scoliosis
6	F	ACLP	—	Delayed psychomotor development
7	M	ACLP	Epicanthus, low-set and malformed ears, short neck	Sacral dimple, clinodactyly of hands, brachydactyly of feet
7A	F	ICL	—	—
8	M	ACL	Hirsutism in front, lower back hairline, low-set and malformed ears, auricular malformation, up-slanting palpebral fissures	Scoliosis, brachydactyly of hands, cryptorchidism, cardiac defects
9	F	ACL	Brachycephaly, flat occipital, prominent frontal, hydrocephaly, prominent glabela, low-set ears, ocular hypertelorism, epicanthus, broad nasal bridge, retrognathia	Delayed psychomotor development, wide-spaced nipples, clinodactyly of hands, flat feet
10	M	ACL	Low-set ears, oblique palpebral fissures down, broad nasal bridge, retrognathia	Wide-spaced nipples, ectrodactyly of hands, congenital clubfoot, hypoplasia of toes, hypotonia, bilateral tibial agenesis
11	F	ACL	Low-set ears, epicanthus, oblique palpebral fissures up, asymmetrical nostrils	Clinodactyly of hands, asymmetrical feet, irregular hyper pigmented lesions on the chest, delayed psychomotor development
12	F	ACL	Macrocephaly, flat occipital, prominent frontal, high frontal hairline low-set ears, preauricular pits, bifid nose	—
13	M	ACP	Long face, prominent supraorbitary crests, malar hypoplasia, depressed nasal bridge and bulbous nasal tip, smooth *philtrum,* thin upper lip, arched palate	*Pectus excavatum*, syndactyly of hands and feet, popliteal *pterygium*, *cubitus valgus*
14	M	ASubCP	Malformed ears, long face, ocular hypertelorism, palpebral ptosis, alar hypoplasia, bulbous nasal tip, epicanthus, strabismus, nystagmus, bilateral aniridia	Ventriculoseptal defect, delayed psychomotor development, hydronephrosis
15	F	ACP	Frontal bossing, capillary hemangioma (frontal, nasal and upper lips), depressed nasal bridge, smooth *philtrum*, retrognathia, low-set and malformed ears	Upper and lower reduction limb defects agenesis of the thumb, arthrogryposis in elbows and knees, partial syndactyly between the 4th and 5th toes, hypoplasia of 5th finger with clinodactyly, Mongolian spots, hypoplasia of *labia majora*
16	M	ICLP	—	—
16U	M	ICLP	—	—
17	M	ICL and SubCP	—	—
18	M	ICLP	—	—
18M	F	ICLP	—	—
19	M	ICL	—	—
20	M	ICL	—	—
21	M	ICL	—	—
22	F	ICL	—	—
23	F	ICP	—	—
23M	F	ICP	—	—

Abbreviations: ACL, associated cleft lip; ACLP, associated cleft lip and palate; ACP, associated cleft palate; ASubCP, associated submucous cleft palate; F, female; ICL, isolated cleft lip; ICLP, isolated cleft lip and palate; ICP, isolated cleft palate; M, male; SubCP, submucous cleft palate.
ID: 7A, aunt of patient 7; 16U, uncle of patient 16; 18M, mother of patient 18; 23M, mother of patient 23; gray color indicates familiar cases.

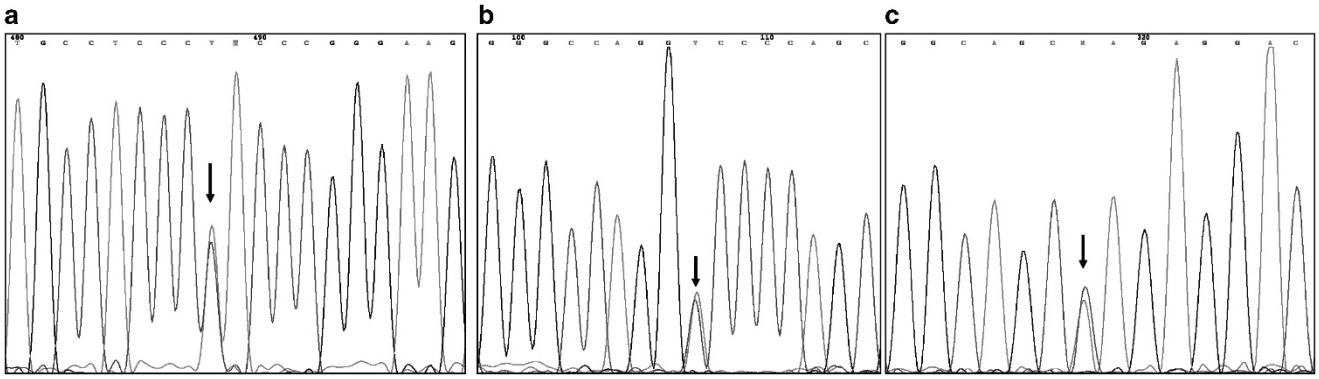

Figure 1 (**a**) DNA sequence of *GLI2* of patient 21. Heterozygous C>T variant sequence is shown by arrow. (**b**) DNA sequence of *MSX1* of patient 8. Heterozygous C>T variant sequence is shown by arrow. (**c**) DNA sequence of *FGF8* of patient 6. Heterozygous C>A variant sequence is shown by arrow. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

FGF18 and *NUDT6*) in a study population consisting of patients with isolated OFC has detected 9 potential pathogenic mutations including one loss-of-function mutation involving *FGF8*. These FGFs genes might be potentially responsible for 3–5% of the isolated CLP cases.[12] Patient 6 uniquely presented CLP and delayed psychomotor development that also corresponded to the timing of *FGF8* expression during the brain development.

Mutations in *GLI2*, *MSX1* and *FGF8* have been reported in single cases of OFC and are therefore considered private mutations.[7] Although this finding might also be detected in the present study sample, our results reinforce the importance of analyzing these genes in patients with OFC. Mutations in genes other than those described by Vieira,[7] which were not investigated in the present study, might be present in other patients of this Brazilian population. In future studies, exome sequencing might identify all genes associated with this disorder.

CNVs have received significant attention in recent years because of the improvements in the resolution of the aGH technique, facilitating its identification in the human genome. CNVs occurring at high frequencies in human populations are considered a potential source of genetic diversity[13,14] and might also be relevant to the pathogenesis of complex traits.[33] The following features of CNVs should therefore be analyzed: the frequency of CNVs in the normal population; the pattern of emergence of CNVs, genes involved with CNVs including patterns of expression and correlated phenotypes; and abnormal phenotypes that have been reported to be associated with the particular CNV.[34–38] When possible, each CNV was analyzed to check whether it was *de novo* or inherited from an affected parent(s) or a normal parent. At the present time, it is not certain whether *de novo* CNV can explain patient's phenotype. The penetrance and extent of the phenotypic spectrum of the imbalance should be considered.[39]

Based on these criteria, CNV analysis was therefore conducted according to three types of references due to the absence of a Brazilian control population database. Considering a minimum size (300 kb) and the number of markers, the analysis generated different results, particularly those derived from the HapMap control data set that mainly assesses genetic differences among various human populations (Supplementary Table 1). Furthermore, the detected CNVs were not recurrent among patients and involved different regions of the genome, indicating the expected OFC genetic heterogeneity based on sample composition.

However, common CNVs among analysis of the same patient were detected (Table 2), except for patients 3, 4, 12, 20 and 22. Most of these have been reported in Database of Genomic Variants, although the population background should also be taken into account. CNV inheritance was predicted in cases in which parental DNA was available, as well as in familial cases. A deletion at the chromosomal region 17q21.31 was detected in patient 7 as well as in his aunt, who was affected by an isolated CLP, highlighting the role of this region that encompassed *ARL17*, *LRRC37A* and *KANSL1*. Haploinsufficiency of *KANSL1* was suggested to cause 17q21.31 microdeletion syndrome (Online Mendelian Inheritance in Man (OMIM): 612452), a multisystem disorder characterized by intellectual disability, hypotonia and distinctive facial features.[40,41] Patient 7 showed a different clinical picture that might probably be related to the size of the deletion.

The most plausible explanation for the majority of detrimental phenotypes caused by changes in copy number is gene dosage, wherein gain or loss of a gene copy causes alteration in expression level.[42] CNVs typically affect multiple genes and, thus, the central question is to estimate the contribution of each gene in CNVs to a particular phenotype.[33] Based on the observed expression pattern and clinical phenotype, we infer that the deletion of the chromosomal region 15q11.2 in patient 17 pointed out *NIPA1*. Mutations in this gene cause hereditary spastic paraplegia type 6, a neurodegenerative disease, and deletions in this gene have been associated with a higher susceptibility to amyotrophic lateral sclerosis.[43,44] *NIPA1* is an inhibitor of *BMP* signaling,[45] and *BMP* genes play a major role in the lip and palate development.[46]

Two CNVs of sizes <300 kb were considered relevant based on the implicated genes: a duplication in patient 3 involved *FGFR1* (chromosomal region 8p12) and a deletion in patient 8 involved *TCEB3* (chromosomal region 1p36.11). *FGFR1* plays a role in palatogenesis,[47] has been linked to OFC development,[48] and mutations in this gene have been implicated in isolated CLP cases.[12] *TCEB3* (transcription elongation factor B (SIII), polypeptide 3) encodes the protein Elongin A, which is a subunit of the transcription factor B (SIII) complex, and is composed of Elongins A/A2, B and C. SIII activates the elongation role of RNA polymerase II by suppressing the transient pausing of the polymerase at several sites within transcription units. Elongin A functions as a transcriptionally active component of SIII complex, whereas Elongins B and C are regulatory subunits.[49] In another study conducted by our group using SNP analysis, this gene was associated with isolated CLP cases (TK Araujo, 2014, unpublished data). In

Table 2 Clinical information (cleft type) and results of genetic analyses including CNV and mutational screening

ID	*Cleft type*	*CNV*	*Range (hg18)*	*Size (kb)*	*Inherited*	*Genes*	*DECIPHER*	*Sequence variants*
1	ACLP	Del 8p23.1	7202009–7847304	645	Dn?	*SPAG11B*, *DEFB*	254627, 251426, 249004	—
1	ACLP	Del 15q11.2	19458525–20089383	631	Dn?	*OR*, *REREP3*	251332, 253672, 249384	—
1	ACLP	Del 16p11.2	32716080–33575301	859	Dn?	*TP53TG3*, *IGHV2OR16-5*	—	—
2	ACLP	Dup 15q11.2	19781829–20089383	308	NF	*OR4,REREP*	—	—
3	ACLP	—	—	—	—	—	—	—
4	ACLP	—	—	—	—	—	—	—
5	ACLP	Dup2q13	110061321–110519397	458	NM	*LIMS3*, *MALL*, *NPHP1*, *NCR*	251953, 248451, 249603	—
6	ACLP	Del 8p23.1	7203124–7733659	531	Dn?	*Genes DEFB (108P1,103A,107B, 104B, 106B, 105B),SPAG11, genes FAM (907, 90A22, 90A23, 90A13), OR7E154P*	—	*FGF8* (Glu236Lis)
7	ACLP	Del 17q21.31	41543275–42107467	564	UA	*ARL17*, *LRRC37A*, *KANSL1*	250499, 2111, 4082, 252142, 17q21.31 microdeletion syndrome	—
7A	ICL	Del 17q21.31	41543275–42107467	564	NA	*ARL17*, *LRRC37A*, *KANSL1*	250499, 2111, 4082, 252142	—
8	ACL	Dup 4q	183613233–183987544	374	NF	*ODZ3*	—	*MSX1* (Ala32Val)
8	ACL	Del 7q11.21	64231764–64619667	388	NF	*ZNF92*	1735	*MSX1*
9	ACL	Del 9p11.2	44149779–44764247	614	Dn	*BX088651.9-2*	—	—
10	ACL	Dup 1q21.1	143662326–144101515	439	Dn?	*PDE4DIP*, *SEC22B*, *NOTCH2NL*, *NBPF10*	251690, 248950, 248281, 249318	—
10	ACL	Dup 7p11.2	56700304–57269864	570	PNM	*ZNF479*, *GUSBP10*	249741	—
10	ACL	Dup 15q11.2	18463963–20089383	1625	NM	*POTE*, *GOLGA8E*, *BCL8*, *HERC2P3*, *CXADRP2*, *REREP3*, *NF1P1*	—	—
11	ACL	Del 15q13.2	28383858–28723577	340	Dn?	*FAM7A1*, *CHRFAM7A*, *ARHGAP11B*	4566, 249439, 773, 249470, 2057, 1226, 248538	—
11	ACL	Dup 16q23.1#	76525303–77158256	633	NM	*WWOX*, *VAT1L*, *CLEC3A*	—	—
11	ACL	Del 22q11.21	17052873–17386984	334	PNM	*DGCR*, *GGT*, *PRODH*	Velocardiofacial syndrome	—
12	ACL	—	—	—	—	—	—	—
13	ACP	Del 1q21.1	146797048–147496455	699	Dn?	*NBPF15*, *NBPF16*	250461	—
13	ACP	Del 2q11.2	97115724–97517954	402	Dn?	*FAHD2B*, *ANKRD36*	—	—
14	ASubCP	Dup8p23.1	8129435–11934586	3800	Dn	*25* genes, including *SOX7* and *GATA4*	249810	—
14	ASubCP	Dup15q25	81869248–98962477	17000	Dn	70 genes, including *KIF7*	—	—
15	ACP	Dup 14q21.3	46190712–46515394	325	NM	*MDGA2*, *RPL10L*	—	—
16	ICLP	Del 1p36.21	12775335–13109945	335	NF	*HNRNPCL1*, *genes PRAMEF* (1,2,4,5,6,7,8,10)	2483, 250348, 255591, 251601, 248448	—
16	ICLP	Del 15q11.2	18671380–19103564	432	NF	*HERC2P2*	—	—
16U	ICLP	Dup 17q21.31	41700624–42092926	392	Dn	*ARL17*, *LRRC37A*, *NSF*	—	—
17	ICL and SubCP	Del 15q11.2	18862054–21038975	2177	PNM	*POTE*, *TUBGCP5*, *CYFIP1*, *NIPA1*, *NFIP2*, *NIPA2*, *NBEAP1*, *GOLGA8E*, *HERC2P3*, *CXADRP2*, *WHAMML1*, *BCL8*, *REREP3*	251332, 253672, 249384	—
18	ICLP	Del 1q21.1	147017157–147603439	586	Dn	*NBPF16*, *DRD5P2*	250369	—
18	ICLP	Dup 15q11.2	18476957–19923712	1447	PAM	*POTE*, *GOLGA8E*, *HERC2P3*, *CXADRP2*, *BCL8*, *NF1P1*	—	—
18M	ICLP	Dup 1p36.13	16743853–17101281	357	Dn?	*ESPNP*, *MST9*, *CROCCP2*, *NBPF1*, *MST1*	—	—
19	ICL	Del 1p36.33#	51586–615321	564	NM	*OR4F5*, *OR4F20*, *OR4F3*, *OR4F29*	2246, 1p36 microdeletion syndrome	—
20	ICL	—	—	—	—	—	—	—
21	ICL	Dup1q21.1	147303136–147689071	386	NM	*HIST2H3PS2, FCGR1C*	248950, 248281, 252452	*GLI2* (Leu761Phe)
21	ICL	Dup 9p11.2	44195870–44840522	645	Dn	*KGFLP1*	—	*GLI2*
22	ICL	—	—	—	—	—	—	—
23	ICP	Dup1q21.1	147303136–147999237	696	AM	*PPIAL4, HIST2H3PS2, CGR1C*	248950, 248281, 252452	—
23	ICP	Del2q13	110218087–110519397	301	NF	*MALL*, *NPHP1*	2111, 251953, 248451	—
23M	ICP	—	—	—	—	—	—	—

Abbreviations: ACL, associated cleft lip; ACLP, associated cleft lip and palate; ACP, associated cleft palate; AM, affected mother; ASubCP, associated submucous cleft palate; CNV, copy number variation; DECIPHER, Database of Chromosomal Imbalance and Phenotype in Humans; Dn, *de novo*; Dn?, suspected of *de novo* although DNA of one parent was not available; F, female; ICL, isolated cleft lip; ICLP, isolated cleft lip and palate; ICP, isolated cleft palate; M, male; NA, nephew affect; NF, normal father; NM, normal mother; PAM, part from affected mother; PNM, part from normal mother; SubCP, submucous cleft palate; UA, uncle affect.
ID: 7A, aunt of patient 7; 16U, uncle of patient 16; 18M, mother of patient 18; 23M, mother of patient 23. CNV: number not reported at Database of Genomic Variants (DGV). Patients 3, 4, 12, 20 and 22 did not present any relevant CNV.

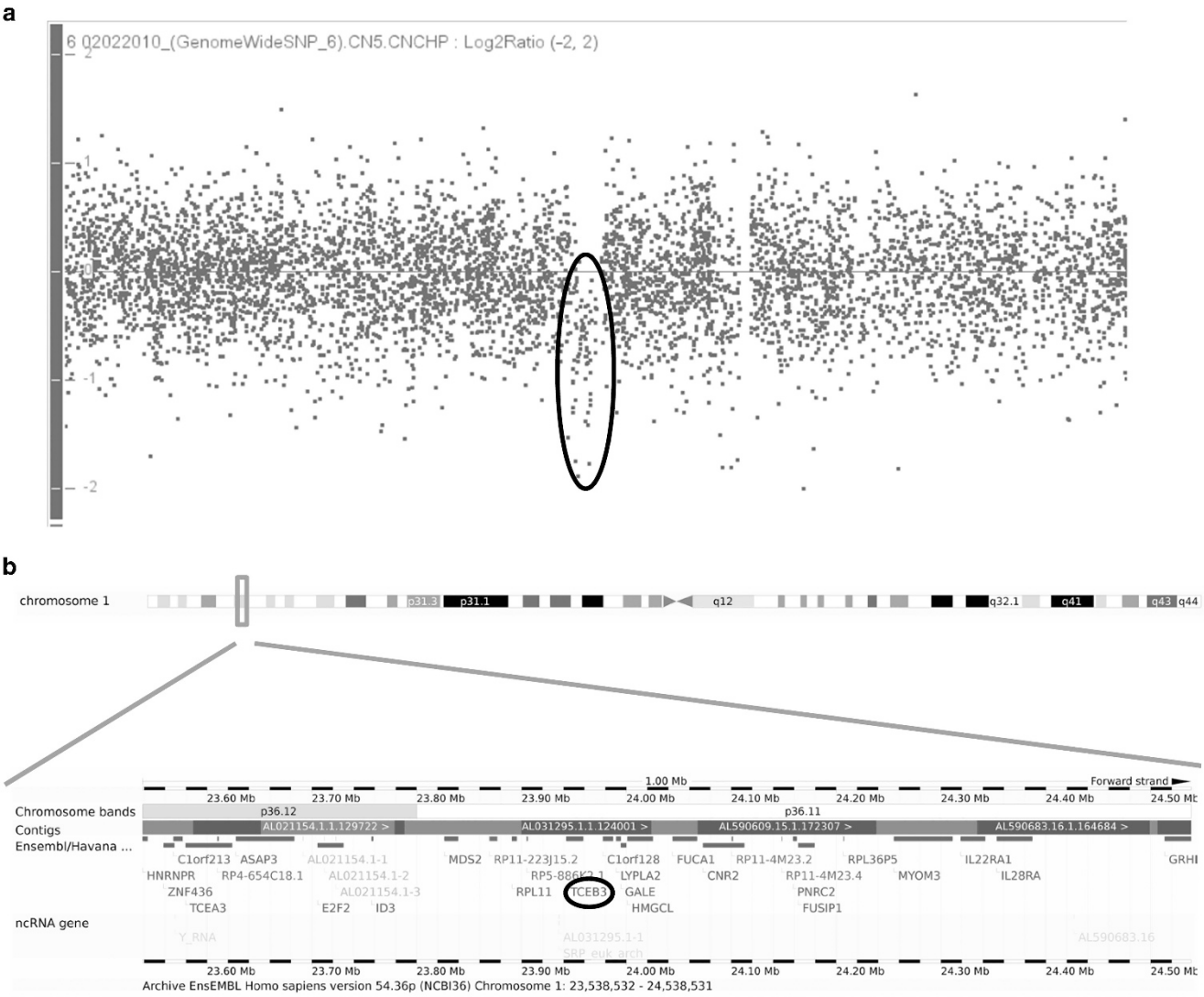

Figure 2 (**a**) Genomic array profile of chromosome 1. Black circle evidences the deletion on 1p36.11 band. (**b**) The region involves *TCEB3* gene (black circle) as showed at Ensembl Genome Browser. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

addition, a 1.88-Mb deletion involving *TCEB3* was detected in case 4661 of the DECIPHER database that presented CP associated with other anomalies.

A 17.09-Mb duplication involving the chromosomal region 15q25–q26 was detected in patient 14. The 15q26 region has been previously reported to be strongly associated with isolated CLP cases in Indian families.[50] In addition, this region encompasses *KIF7* (*kinesin family member 7*) that encodes a protein involved in sonic hedgehog (*SHH*) signaling pathway through the regulation of GLI transcription factors.[51–53] *SHH* signaling plays a role in craniofacial development and lip fusion.[54] Mutations in *KIF7* have been associated with acrocallosal syndrome[53,55,56] as well as hydrolethalus and acrocallosal syndrome, with overlapping features of polydactyly, brain abnormalities and CP.[53] Our group has also associated this gene with isolated CLP cases using SNP analysis (TK Araujo, 2014, unpublished data). Based on all this information, we therefore infer that *KIF7* gene may be suggested to be involved in the development of the submucous CP observed in patient 14.

Considering the phenotype related to CNV, patient 11 harbored a 334-kb deletion that overlapped with the 22q11.2 deletion syndrome (OMIM: #611867) that has also been related to OFC, particularly CP. This region encompasses *PRODH* that is widely expressed in the brain and is a proline dehydrogenase that encodes for the enzyme proline oxidase. This enzyme is responsible for converting proline into glutamate, the main excitatory neurotransmitter of the brain.[57] This gene has been associated with schizophrenia phenotype in the 22q11.2 deletion syndrome.[58–60]

Patient 19, who was classified as an isolated CL case, harbored a 564-kb deletion that overlapped with the 1p36 microdeletion syndrome (OMIM: 607872) region. Monosomy 1p36 is a common terminal deletion syndrome with an estimated incidence rate of ~1 in 5000 births;[61] its main clinical features include developmental delay with hypotonia (100%), seizures (up to 72%), cardiac defects (40%) and CLP (20–40%). The size of the deletion widely varies from ~1.5 Mb to >10 Mb; 40% of the breakpoints occur 3.0–5.0 Mb from the telomere and 70% involve true terminal deletions. A few (~7%) of the 1p36 deletions are interstitial, retaining the 1p subtelomeric region.

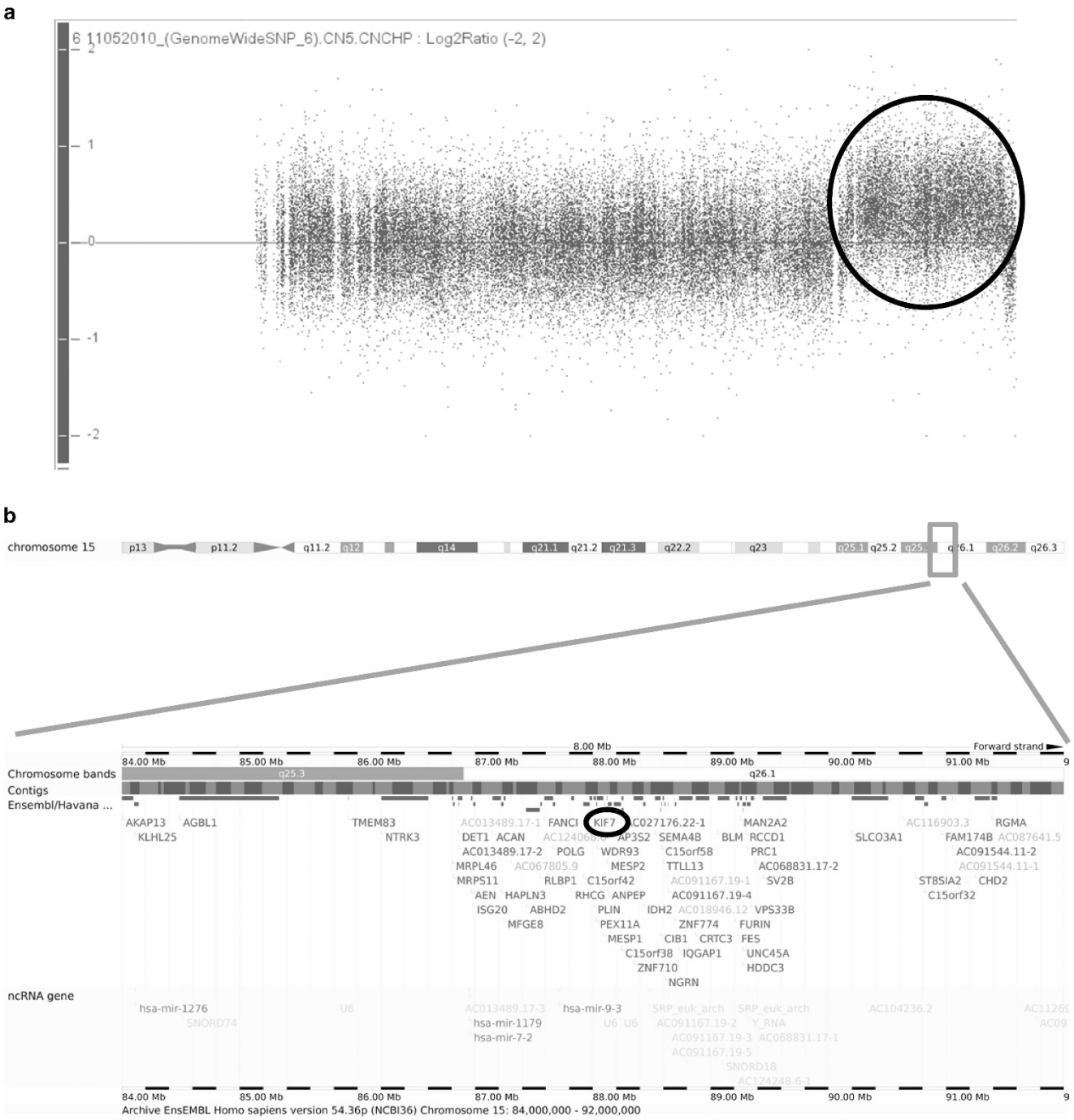

Figure 3 (**a**) Genomic array profile of chromosome 15. Black circle evidences the duplication on 15q25 band. (**b**) The region involves *KIF7* gene (black circle) as showed at Ensembl Genome Browser. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

Breakpoints most commonly occur within the 1p36.13–1p36.33 region that has also been associated by a meta-analysis study to isolated CLP cases,[62] and this is interesting as patient 19 presents isolated CL.

Although extensive studies on CNVs have been conducted, its recent discovery still generates difficulties in its classification and interpretation. For these reasons, understanding the impact of changes in the copy number of individual genes or large chromosomal regions on diseases or malformations is necessary. This study, including a detailed clinical description, contributes to the elucidation of the role of CNVs in OFC pathogenesis. In fact, *major* and *minor* clinical features have been considered relevant to clinical trial.[63]

Considering the complex etiology of OFC, it has been earlier proposed that genetic variations in more than one gene cause this particular phenotype.[7] In addition, the cumulative effect of changes in the copy number of various genes, which individually have little or no effect, could be responsible for the observed abnormal phenotypes.[41] CNV screening has identified new genes that might have influenced OFC pathogenesis and could still be further analyzed. The results of

the present study suggest that a mechanism underlying CNVs associated with sequence variants may play a role in the etiology of this complex congenital defect.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the family cooperation and the Microarray Laboratory of Brazilian Synchrotron Light Laboratory (LNLS) facilities. This report was supported by the Fundação de Amparo à Pesquisa (FAPESP) process number 2008/10596-0. VLGSL is supported by (CNPq 304455/2012-1).

1 Marazita, M. L. The evolution of human genetic studies of cleft lip and cleft palate. *Annu. Rev. Genomics. Hum. Genet.* **13,** 263–283 (2012).
2 IPDTOC Working Group. Prevalence at birth of cleft lip with or without cleft palate: data from the International Perinatal Database of TypicalOral Clefts (IPDTOC). *Cleft Palate Cranio. J.* **48,** 66–81 (2011).
3 Tanaka, S. A., Mahabir, R. C., Jupiter, D. C., Menezes, J. M. Updating the epidemiology of isolated cleft palate. *Plast. Reconstr. Surg.* **131,** 650e–652e (2013).
4 Rittler, M., Cosentino, V., López-Camelo, J. S., Murray, J. C., Wehby, G., Castilla, E. E. Associated anomalies among infants with oral clefts at birth and during a 1-year follow-up. *Am. J. Med. Genet. A* **155A,** 1588–1596 (2011).
5 Marazita, M. L., Mooney, M. P. Current concepts in the embryology and genetics of cleft lip and cleft palate. *Clin. Plastic. Surg.* **31,** 125–140 (2004).
6 Maarse, W., Rozendaal, A. M., Pajkrt, E., Vermeij-Keers, C., Mink van der Molen, A. B., van den Boogaard, M. J. A systematic review of associated structural and chromosomal defects in oral clefts: when is prenatal genetic analysis indicated? *J. Med. Genet.* **49,** 490–498 (2012).
7 Vieira, A. R. Unraveling human cleft lip and palate research. *J. Dent. Res.* **87,** 119–125 (2008).
8 Zucchero, T. M., Cooper, M. E., Maher, B. S., Daack-Hirsch, S., Nepomuceno, B., Ribeiro, L. *et al.* Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *N. Engl. J. Med.* **351,** 769–780 (2004).
9 Vieira, A. R., Ávila, J. R., Daack-Hirsch, S., Dragan, E., Felix, T. M., Rahimov, F. *et al.* Medical sequencing of candidate genes for nonsyndromic cleft lip and palate. *PLoS Genet.* **1,** e64 (2005).
10 Jezewski, P. A., Vieira, A. R., Nishimura, C., Ludwig, B., Johnson, M., O'Brien, S. E. *et al.* Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. *J. Med. Genet.* **40,** 399–407 (2003).
11 Tongkobpetch, S., Siriwan, P., Shotelersuk, V. MSX1 mutations contribute to nonsyndromic cleft lip in a Thai population. *J. Hum. Genet.* **51,** 671–676 (2006).
12 Riley, B. M., Mansilla, M. A., Ma, J., Daack-Hirsch, S., Maher, B. S., Raffensperger, L. M. *et al.* Impaired FGF signaling contributes to cleft lip and palate. *Proc. Natl Acad. Sci. USA* **104,** 4512–4517 (2007).
13 Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P. *et al.* Large-scale copy number polymorphism in the human genome. *Science* **305,** 525–528 (2004).
14 Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y. *et al.* Detection of large-scale variation in the human genome. *Nat. Genet.* **36,** 949–951 (2004).
15 Feuk, L., Carson, A. R., Scherer, S. W. Structural variation in the human genome. *Nat. Rev. Genet.* **7,** 85–97 (2006).
16 Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D. *et al.* Global variation in copy number in the human genome. *Nature* **444,** 444–454 (2006).
17 Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N. *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315,** 848–853 (2007).
18 Osoegawa, K., Vessere, G. M., Utami, K. H., Mansilla, M. A., Johnson, M. K., Riley, B. M. *et al.* Identification of novel candidate genes associated with cleft lip and palate using array comparative genomic hybridisation. *J. Med. Genet.* **45,** 81–86 (2008).
19 Leal, T., Andrieux, J., Duban-Bedu, B., Bouquillon, S., Brevière, G. M., Delobel, B. Array-CGH detection of a de novo 0.8Mb deletion in 19q13.32 associated with mental retardation, cardiac malformation, cleft lip and palate, hearing loss and multiple dysmorphic features. *Eur. J. Med. Genet.* **52,** 62–66 (2009).
20 Shi, M., Mostowska, A., Jugessur, A., Johnson, M. K., Mansilla, M. A., Christensen, K. *et al.* Identification of microdeletions in candidate genes for cleft lip and/or palate. *Birth Defects Res. A Clin. Mol. Teratol.* **85,** 42–51 (2009).
21 Simioni, M., Vieira, T. P., Sgardioli, I. C., Freitas, E. L., Rosenberg, C., Maurer-Morelli, C. V. *et al.* Insertional translocation of 15q25-q26 into 11p13 and duplication at 8p23.1 characterized by high resolution arrays in a boy with congenital malformations and aniridia. *Am. J. Med. Genet. A* **58A,** 2905–2910 (2012).
22 Ruiz-Altaba, A., Palma, V., Dahmane, N. Hedgehog-Gli signalling and the growth of the brain. *Nat. Rev. Neurosci.* **3,** 24–33 (2002).
23 Lebel, M., Mo, R., Shimamura, K., Hui, C. C. Gli2 and Gli3 play distinct roles in the dorsoventral patterning of the mouse hindbrain. *Dev. Biol.* **302,** 345–355 (2007).
24 Bertolacini, C. D., Ribeiro-Bicudo, L. A., Petrin, A., Richieri-Costa, A., Murray, J. C. Clinical findings in patients with GLI2 mutations–phenotypic variability. *Clin. Genet.* **81,** 70–75 (2012).
25 Bear, K. A., Solomon, B. D., Antonini, S., Arnhold, I. J., França, M. M., Gerkes, E. H. *et al.* Pathogenic mutations in GLI2 cause a specific phenotype that is distinct from holoprosencephaly. *J. Med. Genet.* **51,** 413–418 (2014).
26 Zhang, Z., Song, Y., Zhao, X., Zhang, X., Fermin, C., Chen, Y. Rescue of cleft palate in Msx1-deficient mice by transgenic Bmp4 reveals a network of BMP and Shh signaling in the regulation of mammalian palatogenesis. *Development* **129,** 4135–4146 (2002).
27 Satokata, I., Maas, R. Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nat. Genet.* **6,** 348–356 (1994).
28 Ornitz, D. M., Itoh, N. Fibroblast growth factors. *Genome Biol.* **2**: reviews3005.1–3005.12 (2001).
29 Hu, D., Marcucio, R. S., Helms, J. A. A zone of frontonasal ectoderm regulates patterning and growth in the face. *Development* **130,** 1749–1758 (2003).
30 Richman, J. M., Lee, S. H. About face: signals and genes controlling jaw patterning and identity in vertebrates. *Bioessays* **25,** 554–568 (2003).
31 Theil, T., Dominguez-Frutos, E., Schimmang, T. Differential requirements for Fgf3 and Fgf8 during mouse forebrain development. *Dev. Dyn.* **237,** 3417–3423 (2008).
32 Yaguchi, Y., Yu, T., Ahmed, M. U., Berry, M., Mason, I., Basson, M. A. Fibroblast growth factor (FGF) gene expression in the developing cerebellum suggests multiple roles for FGF signaling during cerebellar morphogenesis and development. *Dev. Dyn.* **238,** 2058–2072 (2009).
33 Golzio, C., Katsanis, N. Genetic architecture of reciprocal CNVs. *Curr. Opin. Genet. Dev.* **23,** 240–248 (2013).
34 Lee, C., Iafrate, A. J., Brothman, A. R. Copy number variations and clinical cytogenetic diagnosis of constitutional disorders. *Nat. Genet.* **39**(7 Suppl), S48–S54 (2007).
35 Cook, E. H. Jr., Scherer, S. W. Copy-number variations associated with neuropsychiatric conditions. *Nature* **455,** 919–923 (2008).
36 Koolen, D. A., Pfundt, R., Leeuw, N., Hehir-Kwa, J. Y., Nillesen, W. M., Neefs, I. *et al.* Genomic microarrays in mental retardation: a practical workflow for diagnostic applications. *Hum. Mutat.* **30,** 283–292 (2009).
37 Ionita-Laza, J., Rogers, A. J., Lange, C., Raby, B. A., Lee, C. Genetic association analysis of copy-number variation (CNV) in human disease pathogenesis. *Genomics* **93,** 22–26 (2009).
38 Buysse, K., Delle Chiaie, B., van Coster, R., Loeys, B., de Paepe, A., Mortier, G. *et al.* Challenges for CNV interpretation in clinical molecular karyotyping: lessons learned from a 1001 sample experience. *Eur. J. Med. Genet.* **52,** 398–403 (2009).
39 Thienpont, B., Mertens, L., de Ravel, T., Eyskens, B., Boshoff, D., Maas, N. *et al.* Submicroscopic chromosomal imbalances detected by array-CGH are a frequent cause of congenital heart defects in selected patients. *Eur. Heart. J.* **28,** 2778–2784 (2007).
40 Koolen, D. A., Kramer, J. M., Neveling, K., Nillesen, W. M., Moore-Barton, H. L., Elmslie, F. V. *et al.* Mutations in the chromatin modifier gene KANSL1 cause the 17q21.31 microdeletion syndrome. *Nat. Genet.* **44,** 639–641 (2012).
41 Zollino, M., Orteschi, D., Murdolo, M., Lattante, S., Battaglia, D., Stefanini, C. *et al.* Mutations in KANSL1 cause the 17q21.31 microdeletion syndrome phenotype. *Nat. Genet.* **44,** 636–638 (2012).
42 Tang, Y. C., Amon, A. Gene copy-number alterations: a cost-benefit analysis. *Cell* **152,** 394–405 (2013).
43 Svenstrup, K., Møller, R. S., Christensen, J., Budtz-Jørgensen, E., Gilling, M., Nielsen, J. E. NIPA1 mutation in complex hereditary spastic paraplegia with epilepsy. *Eur. J. Neurol.* **18,** 1197–1199 (2011).
44 Blauw, H. M., van Rheenen, W., Koppers, M., van Damme, P., Waibel, S., Lemmens, R. *et al.* NIPA1 polyalanine repeat expansions are associated with amyotrophic lateral sclerosis. *Hum. Mol. Gene* **21,** 2497–2502 (2012).
45 Tsang, H. T., Edwards, T. L., Wang, X., Connell, J. W., Davies, R. J., Durrington, H. J. *et al.* The hereditary spastic paraplegia proteins NIPA1, spastin and spartin are inhibitors of mammalian BMP signalling. *Hum. Mol. Genet.* **18,** 3805–3821 (2009).
46 Parada, C., Chai, Y. Roles of BMP signaling pathway in lip and palate development. *Front. Oral. Biol.* **16,** 60–70 (2012).
47 Wang, C., Chang, J. Y., Yang, C., Huang, Y., Liu, J., You, P. *et al.* Type 1 fibroblast growth factor receptor in cranial neural crest cell-derived mesenchyme is required for palatogenesis. *J. Biol. Chem.* **288,** 22174–22283 (2013a).
48 Wang, H., Zhang, T., Wu, T., Hetmanski, J. B., Ruczinski, I., Schwender, H. *et al.* The FGF and FGFR gene family and risk of cleft lip with or without cleft palate. *Cleft Palate Cranio. J.* **50,** 96–103 (2013b).
49 Yamazaki, K., Guo, L., Sugahara, K., Zhang, C., Enzan, H., Nakabeppu, Y. *et al.* Identification and biochemical characterization of a novel transcription elongation factor, Elongin A3. *J. Biol. Chem.* **277,** 26444–26451 (2002).
50 Field, L. L., Ray, A. K., Cooper, M. E., Goldstein, T., Shaw, D. F., Marazita, M. L. Genome scan for loci involved in nonsyndromic cleft lip with or without cleft palate in families from West Bengal, India. *Am. J. Med. Gen.* **130A,** 265–271 (2004).
51 Endoh-Yamagami, S., Evangelista, M., Wilson, D., Wen, X., Theunissen, J. W., Phamluong, K. *et al.* The mammalian Cos2 homolog Kif7 plays an essential role in modulating Hh signal transduction during development. *Curr. Biol.* **19,** 1320–1326 (2009).
52 Klejnot, M., Kozielski, F. Structural insights into human Kif7, a kinesin involved in Hedgehog signalling. *Acta Crystallogr. D Biol. Crystallogr.* **68** (Pt 2), 154–159 (2012).

53 Putoux, A., Thomas, S., Coene, K. L., Davis, E. E., Alanay, Y., Ogur, G. *et al.* KIF7 mutations cause fetal hydrolethalus and acrocallosal syndromes. *Nat. Genet.* **43,** 601–606 (2011).
54 Kurosaka, H., Iulianella, A., Williams, T., Trainor, P. A. Disrupting hedgehog and WNT signaling interactions promotes cleft lip pathogenesis. *J. Clin. Invest.* **124,** 1660–1671 (2014).
55 Ali, B. R., Silhavy, J. L., Akawi, N. A., Gleeson, J. G., Al-Gazali, L. A mutation in KIF7 is responsible for the autosomal recessive syndrome of macrocephaly, multiple epiphyseal dysplasia and distinctive facial appearance. *Orphanet J. Rare. Dis.* **7,** 27 (2012).
56 Walsh, D. M., Shalev, S. A., Simpson, M. A., Morgan, N. V., Gelman-Kohan, Z., Chemke, J. *et al.* Acrocallosal syndrome: identification of a novel KIF7 mutation and evidence for oligogenic inheritance. *Eur. J. Med. Genet.* **56,** 39–42 (2013).
57 Boot, E., van Amelsvoort, T. A. Neuroimaging correlates of 22q11.2 deletion syndrome: implications for schizophrenia research. *Curr. Top. Med. Chem.* **12,** 2303–2313 (2012).
58 Jacquet, H., Raux, G., Thibaut, F., Hecketsweiler, B., Houy, E., Demilly, C. *et al.* PRODH mutations and hyperprolinemia in a subset of schizophrenic patients. *Hum. Mol. Genet.* **11,** 2243–2249 (2002).
59 Bender, H. U., Almashanu, S., Steel, G., Hu, C. A., Lin, W. W., Willis, A. *et al.* Functional consequences of PRODH missense mutations. *Am. J. Hum. Genet.* **76,** 409–420 (2005).
60 Prasad, S. E., Howley, S., Murphy, K. C. Candidate genes and the behavioral phenotype in 22q11.2 deletion syndrome. *Dev. Disabil. Res. Rev* **14,** 26–34 (2008).
61 Shaffer, L. G., Lupski, J. R. Molecular mechanisms for constitutional chromosomal rearrangements in humans. *Annu. Rev. Genet.* **34,** 297–329 (2000).
62 Ludwig, K. U., Mangold, E., Herms, S., Nowak, S., Reutter, H., Paul, A. *et al.* Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. *Nat. Genet.* **44,** 968–971 (2012).
63 Puhó, E. H., Czeizel, A. E., Acs, N., Bánhidy, F. Birth outcomes of cases with unclassified multiple congenital abnormalities and pregnancy complications in their mothers depending on the number of component defects. Population-based case-control study. *Congenit. Anom. (Kyoto)* **48,** 126–136 (2008).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)