NEXT GENERATION SEQUENCING INCREASES PRECISE DIAGNOSES IN NON-SYNDROMIC CLEFT LIP AND/OR PALATE

Bénédicte DEMEER

September 2018

Thesis presented to obtain the degree of
Docteur en Sciences Biomédicales et Pharmaceutiques

Promoters
Professor Miikka VIKKULA
Professor Bernard DEVAUCHELLE

UCL Université catholique de Louvain

UNIVERSITÉ de Picardie Jules Verne

de Duve INSTITUTE

NEXT GENERATION SEQUENCING INCREASES PRECISE DIAGNOSES IN NON-SYNDROMIC CLEFT LIP AND/OR PALATE

Bénédicte DEMEER

Thesis publicly defended on September the 17th 2018
to obtain the degree of

Docteur en Sciences Biomédicales et Pharmaceutiques

Jury members:

Professor Pierre COULIE	**President of the jury**
Professor Jill CLAYTON-SMITH	**External member**
Professor Valérie CORMIER-DAIRE	**External member**
Doctor Nicole REVENCU	**Evaluator**
Professor Miikka VIKKULA	**Promoter**
Professor Bernard DEVAUCHELLE	**Promoter**

Cover illustration : Johnny Inukpuk

CARACTERISATION CLINICO-MOLECULAIRE DES FENTES LABIALES ET/OU PALATINES PAR SEQUENÇAGE HAUT DEBIT

Thèse de Doctorat

présentée à l'Ecole Doctorale en Sciences Technologie et Santé (ED 585)

de l'Université de Picardie Jules Verne

par

Bénédicte DEMEER

pour obtenir le grade de
Docteur en Sciences Biologie et Santé

Soutenue le 17 Septembre 2018 après avis des rapporteurs, devant le jury d'examen:

M. P. COULIE, Professeur	Président du jury
M[me] J. CLAYTON-SMITH, Professeur	Rapporteur
M[me] V. CORMIER-DAIRE, Professeur	Rapporteur
M[me] N. REVENCU, Chargée de recherche	Examinateur
M. B. DEVAUCHELLE, Professeur	Directeur de thèse
M. M. VIKKULA, Professeur	Co-encadrant

Cover illustration : Johnny Inukpuk

JURY

Professor Pierre Coulie, president of the jury
Institut de Duve, Belgium

Doctor Nicole Revencu,
Centre de Génétique humaine, Cliniques universitaires Saint-Luc, Université catholique de Louvain, Belgium

Professor Miikka Vikkula,
Human Molecular Genetics (GEHU), de Duve Institute (DDUV), Université catholique de Louvain, Belgium

Professor Bernard Devauchelle,
Service de chirurgie maxillo-faciale, EA CHIMERE, Université de Picardie Jules Verne, CHU Amiens-Picardie, France

External members

Professor Jill Clayton-Smith,
Manchester Center for Genomic Medicine, Manchester Regional Genetics Service, Genetic Medicine, St Mary's Hospital, England

Professor Valérie Cormier-Daire,
Institut de Recherche Necker Enfants Malades, CHU Paris - Hôpital Necker-Enfants Malades, France

ACKNOWLEDGEMENTS

C'est un honneur et une joie de pouvoir remercier les personnes qui ont contribuées à la réussite de cette thèse. Il me sera difficile d'être exhaustive, que les personnes qui seront oubliées m'en excuse.

Mes pensées vont tout d'abord au Pr B. Devauchelle, codirecteur « français » de cette thèse, pour la confiance qu'il m'a accordée depuis ces nombreuses années. Je le remercie pour son appui scientifique, ses conseils et son soutien tant au niveau de la genèse, du suivi que de la finalisation de ce projet.

Je voudrais également adresser tous mes remerciements au Pr M. Vikkula, codirecteur « belge » pour son accueil au sein du laboratoire de génétique moléculaire humaine -GEHU-, institut de Duve, université catholique de Louvain et pour sa confiance en mes capacités à mener à bien ce travail. Merci également pour les différents échanges professionnels et amicaux qui ont contribués à l'aboutissement de ce travail.

Je voudrais remercier l'ensemble des membres de l'équipe pluridisciplinaire des fentes du CHU Amiens-Picardie. Nos collaborations, notamment via les consultations pluridisciplinaires mensuelles, depuis ces longues années ont permis de poser les bases solides de ce travail de thèse pluri et interdisciplinaire. Un merci tout particulier au Dr C. Gbaguidi, pour avoir partagé sa vision « chirurgicale » des fentes, et avoir continué à tenir à jour le « carnet rouge » des patients à discuter.

Je tiens aussi à remercier l'ensemble des membres de l'équipe du centre labio-palatin des cliniques universitaires saint Luc. Les réunions du mercredi matin ont été pour moi un moment privilégié d'échange. Un merci tout particulier aux Docteurs B. Bayet et N. Revencu de m'avoir fait profiter de leur expertise et de leur expérience; ainsi qu'à mesdames M.Valvekens, A.Courtejoie et N.Wade qui participent activement au recrutement des patients porteurs de fentes. Merci au centre de génétique des cliniques universitaires saint Luc pour leur accueil.

J'exprime ma gratitude à mon comité d'encadrement belge constitué par le Pr C. De Smet, le Pr J. B. Demoulin, le Dr N. Revencu, et dirigé par le Pr P. Coulie, pour leurs remarques constructives et leur aide durant ces 5 années.

I would like to thank Pr J. Clayton-Smith and Pr V. Cormier-Daire for granting me the honor to be part of my jury, agreeing to review this manuscript. It is a pleasure to have you for both defences, in Brussels and Amiens.

Je tiens également à remercier le Pr J. P. Marolleau, directeur de l'EA4666, équipe d'accueil au sein de laquelle j'ai débuté ce travail, les Pr S. Testelin et Dr S. Dakpe de la jeune équipe d'accueil CHIMERE/EA7516.

Je tiens aussi à remercier l'ensemble des techniciens, doctorants, post-doctorants et chefs d'équipe du GEHU pour leurs précieux conseils.

Merci au centre de référence maladies rares « anomalies du développement et syndromes malformatifs » Nord de France et à la DRCI du CHU Amiens-Picardie, ainsi qu'aux attachés de recherche clinique et membres du service de chirurgie maxillo-faciale du CHU Amiens-Picardie pour leur soutien.

Je remercie Mesdames M.P. Damonneville et L. Niculescu pour leur soutien administratif, ainsi que les services de scolarité de l'UCL et de l'Ecole Doctorale Sciences Technologie et Santé (ED 585) pour leur aide dans le montage administratif de ce projet Européen.

Je remercie le Pr T. de Ravel de l'Argentière pour ses conseils et son aide.

J'adresse mes remerciements aux patients qui ont acceptés d'intégrer ce projet de recherche.

Je remercie la Région Picardie, devenue Région Hauts de France, pour son soutien financier dans le cadre du projet « GENEPIC ».

Une pensée aux Dr M. Mathieu-Dramard et Dr G. Morin pour leurs implications dans la mise en œuvre de ce projet.

Merci aux amis «belges », natifs ou d'adoption, pour leur présence et soutien.

Enfin, mes derniers remerciements iront à mes parents, pour leur inconditionnel soutien, à mes frères, ma belle-sœur, aux amis proches (CB, BS, IL, JMB, QD, ED, MQ) pour avoir partagé les moments de doute et pour leur présence depuis toutes ces années.

J'ai appris énormément à vos côtés et en suis reconnaissante.

Bénédicte

" Mieux comprendre pour mieux guérir "
Christian de Duve

TABLE OF CONTENT

ABBREVIATIONS

ADH1C: Alcohol dehydrogenase 1c
BMP: Bone morphogenetic proteins
cDNA: complementary DNA
CNCC: Cranial neural crest cells
CL: Cleft lip
CLP: Cleft lip and/or palate
CL/P: Cleft lip with or without palate
CNV: Copy number variation
CP: Cleft palate
CPX: X-linked cleft palate
E: Embryonic day
ECM: Extracellular matrix
EMT: Epithelial Mesenchymal Transformation
FGF: Fibroblast Growth Factor
FGFR: Fibroblast Growth Factor Receptor
FNP: Frontonasal prominence
GAS1: Growth-arrest specific 1
GWAS: Genome-wide association studies
HGP: Human Genome Project
IRF6: Interferon regulatory 6
LNP: Lateral nasal process
LRP (6): Lipoprotein receptor-related protein (6)
MdP: Mandibular prominences
MEE: Medial edge epithelial
MEE: Medial epithelial edge
MES: Medial epithelial seam
MMP3: Matrix metalloproteinase 3
MNP: Medial nasal process
MxP: Maxillary prominence
NS: non-syndromic
NLF: Nasolabial fold
OMIM: Online Mendelian Inheritance in Man
OO(M): Orbicularis oris (muscle)
OFC: Orofacial cleft
OR: Odd ratio

Pc: Post conception
PDGF: Platelet-derived growth factor
PCP: Planar cell polarity
PGDM: Prepregnancy diabetes mellitus
PA: Pharyngeal arches
P63: Tumor protein p63
PRS: Pierre-Robin sequence
PTCH: Patched homolog
PTCH1: Patched homolog 1
PS: Palatal shelves
RNA: Ribonucleic acid
ROR2: Receptor tyrosine kinase-like orphan 2
SMCP: submucous cleft palate
SNP: Single nucleotide polymorphism
mRNA: Messenger RNA
TBX22: T-box 22
TGF: Transforming Growth Factor
TGFβ: Transforming Growth Factor β
SMO: Smoothened
SHH: Sonic Hedgehogh
VPI: Velopharyngeal incompetence
VWS: Van der Woude syndrome
WES: Whole exome sequencing
WGS: Whole genome sequencing
Wnt: Wingless-type MMTV integration site family

SUMMARY

Cleft of the lip and/or palate (CLP) are among the most common birth defects occurring in 1/700 live births. Affected children need multidisciplinary treatment and care from birth until adulthood. CLPs are broadly classified into cleft of the lip with or without palate and cleft of the palate, based on the defective anatomic structures and show a range of phenotypic expression. They can be part of a syndrome, associating additional developmental, structural and/or cognitive anomalies, but most commonly occur as an isolated defect. Non-syndromic (ns) CLPs, of which 20% are familial cases, have a complex etiology with genetic predisposition as well as environmental factors acting in concert. Given the complex and heterogeneous nature of nsCLPs, many genetic approaches have been employed to identify candidate regions and genes for nsCLP, but it is estimated that only 20-25% of the heritability is explained by the different identified regions and loci. With the rapid improvement in technology, in particular invent of next-generation sequencing, from now on we have the capacity to investigate the entire exome or genome in order to identify risk loci. Whole exome sequencing (WES) have now become a useful approach for proving clinical molecular diagnosis.

We conducted WES in a total of 79 nsCLP multiplex families to identify rare causing variants. Among families studied, seven were found to be mutated (~10%) in syndromic genes: *TP63, TBX1, LRP6, GRHL3, TBX22* providing further evidence of contribution of syndromic CLP genes in non-syndromic forms.
Unravelling *TBX22* mutations in 2 families allowed us to describe a new clinical oral sign: bony overgrowths on the posterior edge of the hard palate, on each side of the palatal midline, as well as to expand the *TBX22*-related phenotypic spectrum to choanal atresia and Pierre-Robin sequence.
We also described mutations in *LRRC1* gene in 2 distinct families: one with Van der Woude syndrome and one with ns cleft palate. *LRRC1* had not been implicated in CLP before.

As broad classification used in CLP can overlooked some specific entities, we focused on a distinguishable CLP phenotype. WES was performed on eight sporadic or familial patients with “discontinuous cleft” which combined a cleft of the primary and the secondary palate with preservation of part of the intermaxillary palate. Three causative mutations were found in genes previously implicated in human CLP phenotype and/or syndrome.

SUMMARY

Our work illustrates how well-defined phenotype combined with the power of high throughput technologies significantly contribute to the understanding of genetic factor underlying CLP, and how it could help more accurate clinical, epidemiological, and fundamental research, ultimately resulting in better diagnosis and care of CLP patients.

RESUME

Les fentes labiales et/ou palatines (FLP) représentent les plus fréquentes des malformations congénitales faciales. Elles requièrent une prise en charge multidisciplinaire au long terme. Elles se produisent majoritairement de façon isolée. Leur cause est complexe et hétérogène, associant des interactions entre prédisposition génétique et facteurs environnementaux. Parmi les nouvelles technologies d'étude du génome, le séquençage haut débit d'exome (WES) est devenu un outil diagnostique moléculaire majeur.

Nous avons étudié par WES 79 familles multiplexes porteuses de FLP non-syndromique pour identifier des variants rares causaux.
Chez sept d'entre elles, soit chez environ 10%, une mutation dans un gène impliqué dans les FLPs syndromiques ont été mises en évidence : *TP63, TBX1, LRP6, GRHL3* et *TBX22* renforçant la notion de contribution de certains gènes communs aux formes syndromiques et non syndromiques.
La réévaluation clinique des 2 familles porteuses de mutation dans *TBX22* a permis de décrire un nouveau signe clinique, sous la forme d'excroissances osseuses situées à la partie postérieure du palais dur, de chaque côté de la ligne médiane palatine; et d'élargir le spectre clinique associé avec description d'atrésie choanale et séquence de Pierre-Robin.
Nous avons également trouvé des mutations dans un nouveau gène, *LRRC1*, chez 2 familles distinctes.

De plus nous avons étudié par WES 8 cas sporadiques ou familiaux porteurs d'un phénotype reconnaissable nommé «FLP discontinue», défini par l'association d'une fente du palais primaire et du palais secondaire avec préservation d'une partie du palais intermaxillaire. Trois mutations (>1/3) ont été trouvées dans des gènes précédemment impliqués dans les FLPs.

Notre travail illustre qu'un phénotypage précis combiné à des méthodes d'investigations génétiques à haut-débit, contribuent de manière significative à une meilleure compréhension des facteurs génétiques des FLPs, ainsi qu'à la recherche clinique, épidémiologique et fondamentale, dans le but ultime d'améliorer le diagnostic et la prise en charge des patients.

RESUME

Les fentes labiales et/ou palatines (FLPs) représentent les plus fréquentes des malformations cranio-faciales congénitales avec une incidence estimée à 1/700 naissances, incidence présentant des variations selon les ethnies et le type de fentes. Elles sont reparties en fente labiale avec ou sans fente palatine (FL/P) et en fente palatine seule (FP). En accord avec les différentes études épidémiologiques et génétiques, elles sont majoritairement non syndromiques -70% pour les FLP, 50% pour les FP-, c'est à dire sans autre anomalie associée. Dans nos pays occidentaux les fentes labiales et/ou palatines n'entraînent pas d'augmentation de la mortalité, elles nécessitent cependant un traitement multidisciplinaire chronique et coûteux. Une origine multifactorielle est le plus souvent retenue dans la genèse des FLPs, avec une interaction complexe entre facteurs environnementaux et facteurs génétiques. La qualité descriptive des phénotypes et la classification des associations malformatives en sous-groupes a largement contribué à l'identification des gènes responsables de FLPs notamment syndromiques chez l'homme. De plus, et pour tenir compte des problèmes liés à leur hétérogénéité, des études basées sur une description clinique précise et de « sous-phénotypes » associés semblent prometteuses. Le spectre phénotypique associé aux FLPs non-syndromiques présente une complexité plus importante qu'initialement pensée. En complément des FLPs « manifestes », certaines atteintes s'avèrent « discrètes », plus difficiles à diagnostiquer, et peuvent être notées, tant au niveau labial supérieur (indentations linéaires avec ou sans déformation narinaire, rare fistule de la lèvre supérieure), qu'au niveau palatin (luette bifide, fente sous muqueuse). Enfin, le spectre phénotypique lié aux FLPs s'est enrichi par la description de signes « mineurs », observés chez les individus porteurs de FLPs et également chez les apparentés « non atteints » tels que : de petites dépressions de la lèvre inférieure, des anomalies dentaires, une atteinte du muscle orbicularis oris, une discontinuité du sillon naso-labial, des particularités morphologiques et morphométriques faciales (imagerie biométrique faciale 3D non-invasive), et de possible atteinte vélopharyngée, olfactive ou cérébrale associée.

L'identification de gènes ou de loci chromosomiques impliqués dans les fentes labiales et/ou palatines est le résultat de dizaines d'années de recherche et de l'utilisation de multiples approches génétiques, souvent complémentaires comme : l'identification et l'étude d'anomalie chromosomique ou de microdélétion au sein d'individu atteint, les études familiales par analyse de liaison, le séquençage direct,

les études de gènes candidats et études d'association. L'étude de modèles animaux, et plus particulièrement de la souris, a été à l'origine de la découverte de nombreux gènes candidats, et a aidé à la compréhension de la fonction et l'implication de nombreux gènes dans le développement facial. Une coordination spatio-temporelle précise de différents types cellulaires, cellules mésenchymateuses et épithéliales notamment, est nécessaire à la morphogenèse faciale. De nombreux gènes et voies de signalisation moléculaire telles que les voies *FGF, WNT, TGF, BMP, PDGF* et *NOTCH* ont été impliquées dans cette morphogenèse, ainsi que certains mécanismes génétiques ou épigénétiques et des facteurs environnementaux. Cependant malgré des progrès significatifs dans l'identification de gènes et voies de signalisation impliqués dans le développement craniofacial, ces différentes approches n'ont permis d'expliquer qu'une partie des bases moléculaires des FLPs chez l'homme.

D'importantes découvertes dans les formes non-syndromiques de FLP ont émané de l'étude de formes syndromiques. Certains gènes de FLPs syndromiques ont été impliqués dans les FLPs non-syndromiques, et associés à une pénétrance incomplète peut être secondaire à l'action de facteurs modificateurs. La plus fréquente forme syndromique de FLP est le syndrome de Van der Woude (VWS), causé pour 70 % des cas par une mutation du gène *IRF6* et pour 5% des cas par une mutation du gène *GRHL3*. Une pénétrance incomplète est observée dans le VWS, et il ne se distingue phénotypiquement des formes non-syndromiques de FLP que par des signes orofaciaux « mineurs » comme de petites dépressions de la lèvre inférieure (lip pits) ou une hypodontie, présents de façon variable. Par conséquent 10 à 15 % des patients VWS se présentent comme des formes non syndromiques de FLP. Des mutations dans le gène *MSX1* sont associées aux FLPs syndromiques (avec hypodontie) et dans approximativement 2% des cas aux FLPs non syndromiques. De nombreux autres gènes de formes syndromiques, comme *FGFR1, TP63, TBX22, PVRL1* and *PTCH,* étudiés pour leur rôle dans les formes syndromiques pourraient contribuer aux formes non-syndromiques.

Etant donné la complexité des étiologies associées aux formes non syndromiques de FLP, l'identification de nouveaux gènes causaux représente un réel défi. L'essor des nouvelles technologies, et notamment l'avènement du séquençage à haut débit, marque une nouvelle ère dans laquelle il sera possible, pour un coût raisonnable, de réaliser le séquençage du génome de sujets atteints. Parmi les nouvelles technologies d'étude du génome, le séquençage haut débit d'exome (WES) a démontré sa capacité à détecter des mutations causales pour de nombreux

syndromes. Depuis 2010, son utilisation a permis l'identification croissante de gènes responsables de maladies rares et est devenu un outil diagnostique moléculaire majeur.

Nous avons pu étudier par séquençage haut débit d'exome 79 familles multiplexes porteuses de FLP non-syndromique afin d'identifier des variants rares causaux.

En appliquant une approche « gène candidat », nous avons filtré les variants rares potentiellement causaux obtenus par WES à l'aide d'une liste de gènes candidats. Chez sept d'entre elles (~10%) une mutation dans un gène impliqué dans les FLPs syndromiques a été mise en évidence : *TP63, TBX1, LRP6, GRHL3* et *TBX22* renforçant la notion de contribution de certains gènes communs aux formes syndromiques et non syndromiques.
Concernant les familles porteuses de mutations au sein des gènes *TP63, TBX1, LRP6* et *GRHL3*, la réévaluation clinique des patients a permis de confirmer le caractère non syndromique de la FLP et de souligner l'extrême variabilité d'expression phénotypique associée s'étendant de forme complète de FLP à de discrète atteinte velo-pharyngée. Il est intéressant de noter que pour ces gènes, lorsque les mutations sont classées en mutation perte de fonction ou mutation faux-sens, les mutations décrites appartiennent au groupe le moins fréquent pour un gène donné.
Pour les gènes *TP63* et *TBX1*, les rares mutations « perte de fonction » connues sont regroupées au niveau de la partie 3' du gène et sont classiquement associées à un phénotype syndromique. Les 2 mutations décrites dans notre étude sont localisées au niveau du domaine de liaison à l'ADN, situées au milieu des protéines TP63 et TBX1. Ainsi pour ces gènes, une certaine corrélation génotype-phénotype concernant le type et/ou la localisation des mutations est évoquée.
La mutation c.1171C>T (NM_198174.2 ; p.Arg391Cys, NP_002327.2) du gène *GRHL3* décrite chez 2 des familles étudiées, possède un effet dominant négatif, et a été associée dans la littérature à un syndrome de Van der Woude de type 2 (VWS2) chez un patient porteur d'une FLP et dépressions labiales inférieures. Sa description dans notre étude chez 2 nouvelles familles permet de confirmer son association à une variabilité phénotypique intra et interfamiliale, ainsi que de l'associer à différentes formes non syndromiques de FLPs comme une fente palatine isolée et une séquence de Pierre-Robin. L'hypothèse d'un point chaud mutationnel est évoqué devant cette transformation C>T, au niveau d'un ilôt CpG, décrite chez des individus d'origine ethnique différente.

Concernant le facteur de transcription *TBX22*, ses mutations ont été associées au syndrome de fente palatine et ankyloglossie lié à l'X -CPX (MIM 303400)-. Le syndrome CPX résulte vraisemblablement, comme d'autres pathologies liées aux gènes de la famille TBX, d'une perte de fonction de la protéine TBX22. Les 2 mutations de *TBX22* décrites dans notre étude sont des mutations canoniques d'un site donneur d'épissage. Dans la forme classique de CPX, il existe une importante variation inter et intra-familiale, et le spectre phénotypique associé peut s'étendre d'une ankyloglossie isolée à une atteinte palatine de type fente sous-muqueuse, luette bifide, voire fente du palais secondaire isolée ou associée à une ankyloglossie. La mutation c.633+1G>A (NM_016954) touchant le site donneur consensus de l'exon 4 a été précédemment mise en évidence au sein de 2 familles porteuses d'une forme classique de syndrome CPX. La mutation touchant le site donneur GT consensuel de l'exon 2 (c.356+1G>A; NM_016954) constitue une nouvelle mutation, et s'associe à une forme classique de CPX, avec variabilité d'expression intrafamiliale et pénétrance incomplète, et chez le cas index à une séquence de Pierre-Robin. Concernant la famille porteuse de la mutation c.633+1G>A sa réévaluation clinique a permis de décrire un nouveau signe clinique sous la forme d'excroissances osseuses situées à la partie postérieure du palais dur, de chaque côté de la ligne médiane palatine. Il a été montré chez la souris KO *Tbx22* que Tbx22 possède un rôle majeur dans la formation osseuse intramembraneuse du palais osseux postérieur, plutôt que d'être impliqué dans la fusion des lames palatines. Ainsi la description de ces excroissances osseuses et leur localisation posent la question d'une anomalie (réactionnelle) d'ossification de la partie postérieure palatine associée. Des descriptions complémentaires sont nécessaires pour conclure quant à la spécificité de ce signe clinique. Un phénotypage précis de cette famille a également permis d'élargir le spectre clinique associé au gène *TBX22* par la description d'une atrésie choanale chez certains patients porteurs de la mutation. Il est intéressant de noter que le modèle souris KO *tbx22* présente des malformations craniofaciales variables et multiples, compatibles avec le modèle d'expression connu de *tbx22*, comme une fente sous-muqueuse et pour une minorité, une fente palatine associée ou non à une réduction de taille du vomer et une ankyloglossie. Certaines souris *tbx22 null* sont porteuses également d'une atrésie choanale. Des études complémentaires sont nécessaires pour déterminer la variabilité et la pénétrance de l'atrésie choanale dans les pathologies liées à *TBX22*.

L'application d' une approche « gène commun » nous a permis de mettre en évidence des mutations au sein du gène *LRRC1*, non connu à ce jour pour être

impliqué dans le développement craniofacial, au sein de 2 familles distinctes: la première famille présentant un syndrome de Van der Woude, la seconde une fente palatine non-syndromique. Les 2 mutations sont des mutations non-sens situées en 3' de la protéine pour lesquelles nous avons pu réaliser des études fonctionnelles. Au niveau de l'ARN messager issu de lymphoblastes d'individus porteurs de ces mutations, le séquençage « Sanger » du cDNA a permis de montrer la présence en quantité égale du cDNA « sauvage » et « muté ». Au niveau protéique, l'étude de l'expression de LRRC1 dans des cellules Cos7 exprimant la protéine LRRC1 sauvage et mutée a permis de montrer l'expression d'une protéine mutée de taille réduite, en quantité conservée. Ainsi les 2 mutations identifiées résultent en la production d'une protéine mutée stable, tronquée, ne possédant plus le domaine de liaison PDZ de la protéine situé dans sa partie C-terminale.

Peu de données sont disponibles sur le gène *LRRC1*. *LRRC1* est exprimé de façon ubiquitaire dans les tissus humains adultes. Durant l'embryogenèse, seules quelques données chez la souris sont disponibles avec l'existence d'une expression orofaciale au sein de la région maxillaire apparaissant avant le début de la palatogenèse mais le profil d'expression orofaciale n'est pas documenté après E10,5. Des études d'expression complémentaires sont donc nécessaires, notamment chez l'humain. LRRC1 appartient à la famille des protéines LAP, protéines impliquées dans l'établissement de la polarité des cellules épithéliales, et qui agissent comme « adaptateur » dans différentes voies de signalisation intracellulaire. Pour les mutations décrites, l'hypothèse est celle de protéines LRRC1 mutées stables, qui garderaient leur localisation habituelle latérobasale au sein des cellules épithéliales, mais qui seraient incapables d'interagir avec les domaines PDZ de ses partenaires. L'implication de LRRC1 dans différents complexes protéiques fait de cette protéine un gène candidat d'intérêt dans les FLPs. En effet, un de ses partenaires, *DLG1,* est un gène candidat pour les FLPs non syndromiques. Un autre de ses partenaires, *ERBIN*, possède un rôle inhibiteur de la voie de signalisation TGFB, elle-même impliquée dans la transformation epithélio-mesenchymateuse des cellules épithéliales et dans la fusion palatine. Des études supplémentaires de cohortes, ainsi que la détermination du profil d'expression embryonnaire durant le développement orofacial, et des expériences fonctionnelles complémentaires sont nécessaires pour déterminer le rôle potentiel du gène *LRRC1* dans le développement craniofacial.

Enfin, nous nous sommes intéressés à un phénotype reconnaissable nommé « FLP discontinue », défini par l'association d'une fente du palais primaire et du palais secondaire avec préservation d'une partie du palais intermaxillaire. Bien que ce phénotype puisse représenter 5% des FLPs, il est rarement fait spécifiquement

référence à celui-ci, et sa physiopathologie reste inconnue. Nous avons étudié par WES huit cas sporadiques ou familiaux de FLP non-syndromique discontinue en utilisant une approche « gène candidat ». Nous avons mis en évidence une mutation perte de fonction chez trois individus, ce qui représente près de 40% de cette cohorte de petite taille, et impliqué les gènes *FGFR1, ARHGAP29* et *DLG1*.

Le gène FGFR1 Fibroblast Growth Factor 1, appartient à la voie de signalisation FGF qui possède un rôle essentiel dans la palatogenèse. FGFR1 a précédemment été impliqué dans les formes syndromiques ou non-syndromiques de FLP, parfois au sein d'une même famille, et associé au FL/P et FP.

Le gène *DLG1* représente un gène candidat de FL/P non syndromique. Le mutant souris dlg1 *null* décède rapidement après la naissance, et présente des malformations congénitales majeures avec atteinte cranio-faciale et possible FP. DLG1 est exprimée dans les cellules mésenchymateuses et épithéliales de l'ensemble du palais. De plus, les interactions protéiques impliquant les domaines C-terminaux semblent essentielles pour la normalité de sa fonction dans la morphogénèse cranio-faciale et la palatogenèse, et il est suggéré que la perte de dlg1 puisse affecter le développement craniofacial via la voie de signalisation Wnt/PCP. Enfin, en pathologie humaine, 4 à 9 % des individus porteurs d'une microdélétion interstitielle 3q29, incluant *DLG1*, présentent une FLP.

Il a récemment été conclu que le gène *ARHGAP29* correspond au gène responsable de FLP situé au sein du locus 1p22. Des études de séquençage ont montré que des variations perte de fonction au sein de ce gène contribuaient au FL/P non-syndromique. Plus récemment, une mutation faux-sens a été impliquée au sein d'une forme familiale de FP. Ainsi les variants pathogènes d'*ARHGAP29* sont associés au FLP avec variabilité d'expression inter et intrafamiliale, et au sein des formes familiales, à une transmission de type autosomique dominante avec pénétrance incomplète. Il est à noter que la mutation perte de fonction décrite dans notre étude coségrège dans une forme familiale de FLP, dont le cas index présente une fente discontinue avec atteinte labiale de type « microforme », phénotype labiale déjà décrit au sein d'une forme familiale de FLP avec mutation d'*ARHGAP29*.

La littérature rapporte, au sein de la cohorte clinique de 356 cas de FLP discontinue publiée par Alvarez et coauteurs, quatre patients cliniquement diagnostiqués comme syndrome de Van der Woude -VWS- et un comme syndrome des pterygiums poplités -PPS-. En tenant compte de l'implication du gène *IRF6* chez respectivement 70% des cas de VWS et 97% des cas de PPS, ainsi que dans les FLPs non-syndromiques, le gène *IRF6* apparaît comme un gène candidat d'intérêt dans les FLPs discontinues.

Notre travail illustre qu'un phénotypage précis combiné à des méthodes d'investigations génétiques à haut-débit, contribuent de manière significative à une meilleure compréhension des facteurs génétiques des FLPs, ainsi qu'à la recherche clinique, épidémiologique et fondamentale, dans le but ultime d'améliorer le diagnostic et la prise en charge des patients.

Le séquençage haut débit d'exome représente actuellement le meilleur moyen d'étude des formes familiales de FLP, ainsi que de certains phénotypes reconnaissables.

Du fait de l'évolution majeure des techniques d'étude, les différents systèmes de classification utilisés, majoritairement basés sur l'anatomie et sur l'embryologie, s'avèrent actuellement insuffisants. Dans un but de recherche ces classifications se doivent d'être affinées et enrichies tant en termes de génétique que de phénotypage. Le développement de nouveaux outils de «phénomique» permet par exemple d'envisager de regrouper et de comparer les données cliniques ainsi que de faciliter les corrélations avec le nombre croissant de données génétiques engendrées par les différentes techniques de séquençage haut débit.

Des progrès notables ont été réalisés dans l'identification des gènes impliqués dans les FLPs, différentes voies de signalisation ont été impliquées dans les FLPs non-syndromiques, mais aucune voie biologique n'est à ce jour clairement établie dans la formation des FLPs. En identifiant des mutations au sein de gènes fonctionnellement liés, l'utilisation du WES a récemment permis de mettre en évidence une nouvelle voie biologique importante dans la formation des FL/Ps, voie impliquant le mécanisme d'adhésion épithéliale via le complexe cadhérine-caténine épithélial.

A l'échelle individuelle et de la population, comprendre les facteurs sous-tendant les FLPs non-syndromiques représente un préalable nécessaire à l'amélioration de la prévention, du traitement, du pronostic et des connaissances des facteurs génétiques et environnementaux de ce type d'atteinte.

Au niveau cellulaire et moléculaire, le développement des nouvelles disciplines biologiques de type « -omique » telles que la génomique, l'épigénomique, la transcriptomique, et la protéomique appliquées à différents types cellulaires et tissulaires devraient permettre d'aider à élucider les mécanismes impliqués dans les FLPs.

I. Introduction

A. Craniofacial development

During early embryogenesis, the gastrulation phase gives rise to three primary germ cell layers called ectoderm, mesoderm and endoderm. The neural plate, corresponding to thickening of the ectoderm, folds in upon itself and form the neural tube. At the neural plate border, induction of cells competent to generate cranial neural crest cells (CNCC) occurs. During neural tube closure (in the fourth week of embryogenesis), CNCCs delaminate and migrate as mesenchyme into the developing embryonic processes of the head and neck. This migration is crucial for facial development and combines with the core mesoderm and epithelial cover to set up the 5 facial primordia : the rostral frontonasal prominence (FNP), the two lateral maxillary (MxP) and the two caudal mandibular prominences (MdP) [1]. They give rise to a large diversity of cell and tissue types in the craniofacial region; with rostral CNCCs migrating into the frontonasal process (FNP) forming the forehead, philtrum of the upper lip and nose; and more posterior CNCCs migrating into the four pharyngeal arches (PA), the first PA forming the maxillae and the mandible. On the other hand, facial and lingual muscles derive from the mesoderm [2]. Precisely orchestrated reciprocal signaling interactions between CNCCs and neighboring tissues are needed for a proper morphogenesis with differentiation, growth, and eventual fusion of the five facial prominences that will form the definitive face [2].

1. Lip, alveolar and primary palate formation

As facial prominences bulge, they border a deepened central depression in the facial region, the stomodeum, precursor of the future mouth. From the fourth to fifth week of human embryogenesis, the frontonasal prominence widens and thickening of surface ectoderm occurs bilaterally on the ventrolateral part giving rise to the nasal placodes (24 day post-conception (pc) in human). The latter invaginate and form paired medial (MNP) and lateral nasal (LNP) processes. Rapid growth of MNP and maxillary processes (MxP) wedges the LNP rostrally, in between, and brings the distal ends of maxillary and medial nasal processes into direct contact. Initial fusion is between LNP and MNP. It is followed by fusion between MNP and MxP that gives rise to the intermaxillary segment, which will become the philtrum and the upper lip. Its outgrowth into the oral cavity will form the triangular shaped primary palate extending posteriorly to the incisive foramen

(or clinically to the incisive papilla) and anteriorly including the alveolar ridge with its four incisor teeth. Fusion between these prominences involves active epithelial filopodial that emerge when they are brought into proximity to each other's, to establish bridges, and adhering interactions to maintain them in close proximity [3]. Periderm cells covering will undergo apoptosis, resulting in their slugging off and allowing the opposing basal epithelial cells to form a transient epithelial seam. The latter will disintegrate by apoptosis [3, 4] and/or Epithelial Mesenchymal Transformation (EMT) [5] to allow mesenchymal confluence and complete upper lip formation by 48 days pc in human. It is thought that similar mechanisms are involved in lip fusion and secondary palate fusion [3-5]. Once upper lip morphogenesis is complete, the upper lip is composed of tissues derived from the medial nasal and maxillary processes and the lateral nasal processes form the sides (alae) of the nose (Figure 1). The primary palate will later fuse with the secondary palate derived from the maxillary processes.
Intramembranous ossification of the primary palate begins at the eighth week of human embryonic development at the medial nasal prominences and spreads laterally across the maxillary prominences. Bone differentiation is accompanied by the development of facial musculature.

The paired mandibular processes, which originate from the first pair of pharyngeal arches, will merge to form the lower jaw and lip, first parts of the face to be completed, at day 37 pc in human. The earlier establishment of the lower lip might be one of the reason of the rarity of lower lip clefting [6], as it provides a shorter period for teratogenic agents to disrupt the merging procedure, although genetic factors are likely to be implicated as well.

Deficit in facial process size, misalignment of facial processes, or failure of fusion of the epithelial seams result in a cleft of the lip. The cleft can be uni or bilateral, and of varying degrees of severity: it can comprise the lip only, or extend to the alveolar bone with or without the primary palate (Figures 2 and 3). Although the lateral nasal processes do not contribute to the final upper lip, failure of fusion of the medial nasal processes with both maxillary and lateral nasal processes during upper lip development will give rise to a cleft of the lip extending to the nostril [4]. Minimal lip involvement -which manifests as linear lip indentation or lateral fistulas of the upper lip- is at the position of fusion of the median nasal and maxillary prominences [7].

2. Secondary palate formation

Early in the 6th week of embryogenesis, the secondary palate arises as paired outgrowths from the oral side of the maxillary processes, which initially grow vertically down on both side of the developing tongue. The shelves undergo a rapid change in position and shape to become horizontally oriented toward one another above the tongue. Palatal shelf elevation is driven by "intrinsic force" within the palatal shelves, that is still somewhat incompletely understood and influenced by other craniofacial and oral structures, including growth of the stomodeum and embryo mouth-opening reflexes allowing the tongue to flatten [8]. Several constituents, including extracellular matrix (ECM) components and transcription factors that are differentially expressed in developing palatal tissue, contribute to this intrinsic elevating force and palatal shelf reorientation occurs in a dynamic and regionalized manner along the antero-posterior axis [9, 10]. The progressive regionally mesenchymal specific accumulation of glycosaminoglycans, primarily hyaluronic acid, will result in swelling of the ECM, by retaining high amount of water, could generate the force for palatal elevation and contribute to "reshape" the tissues. In addition, an alignment of F-actin fibers toward the medial wall in the middle and posterior regions of palatal shelves is in favour of an actin-based cell contraction that drives their reorientation [10]. The horizontally positioned palatal shelves grow towards the midline and fuse by the tenth week. Shelf adhesion is temporally and spatially controlled to avoid aberrant adhesion to other structures. It requires a proper regulation of medial epithelial edges (MEE) differentiation and adhesion competence, via junction complexes, which include the adherent junctions (cadherin and nectins), tight junctions, desmosomes and growth factors, and creates a transient medial epithelial seam (MES). MES will disintegrate to finally provide ultimate mesenchymal continuity throughout the fused palate. The fate of the MES/MEE has been subject to considerable debates, and three non-exclusive mechanisms have been proposed: death of the MES, by apoptotic or non-apoptotic mechanisms, epithelial-to-mesenchymal transformation (EMT) and migration of MES cells towards the periphery of the midline [11]. As mesenchyme continuity across the intact horizontal palate establishes, the epithelia on the nasal aspect of the palate differentiate into pseudostratified ciliated columnar cells, whilst those on the oral aspect of the palate become stratified squamous, non-keratinizing cells [8].

The secondary palate fuses anteriorly with the primary palate and anterodorsally with the nasal septum to form the intact roof of the oral cavity. Most of nasopalatine canal is obliterated except for the incisive foramen. Palatogenesis

initiated in the sixth week, is completed by the twelfth week of gestation [9] (Figure 1).
Cranial vault and most of the facial bones, including the palatal bone, form by intramembranous ossification. This process involves condensation of the neural crest derived mesenchyme cells that differentiate directly into osteoblasts. At the end of the seventh week prior to elevation of PS, osteogenic blastemata for the palatal processes of the maxillary and palatine bones differentiate in the mesenchyme of the hard palate [3] from "de novo ossification centers" that are initially separated from the maxillary bone and by expansion of osteogenic fronts from the existing palatine towards the midline respectively [12], to form the anterior bony "hard" palate. In parallel, several patterned myogenic blastemata develop in order to form the posterior part of the palate, the "soft "palate and uvula, the last portion of the palate to form.
The hard palate acts as a static partition between the mouth and the nasal cavities, enabling simultaneous breathing and food intake, while the soft palate acts as a dynamic barrier obstructing nasal passages during swallowing and intermittently valving for speech and airway.

Cleft palate results from disturbances in any stage of palatal development: defective palatal shelf growth, delayed or failed shelf elevation, defective fusion, post-fusion rupture and failure of mesenchymal differentiation [8] (Figure 2 and 3). Defects in the palatal skeleton formation, after proper fusion of the palatal shelves, result in submucous cleft palate (SMCP). Cleft palate can also result from a defect in other craniofacial structures (i.e. mandible or tongue); the most evoked example is clinically referred as Pierre-Robin sequence.

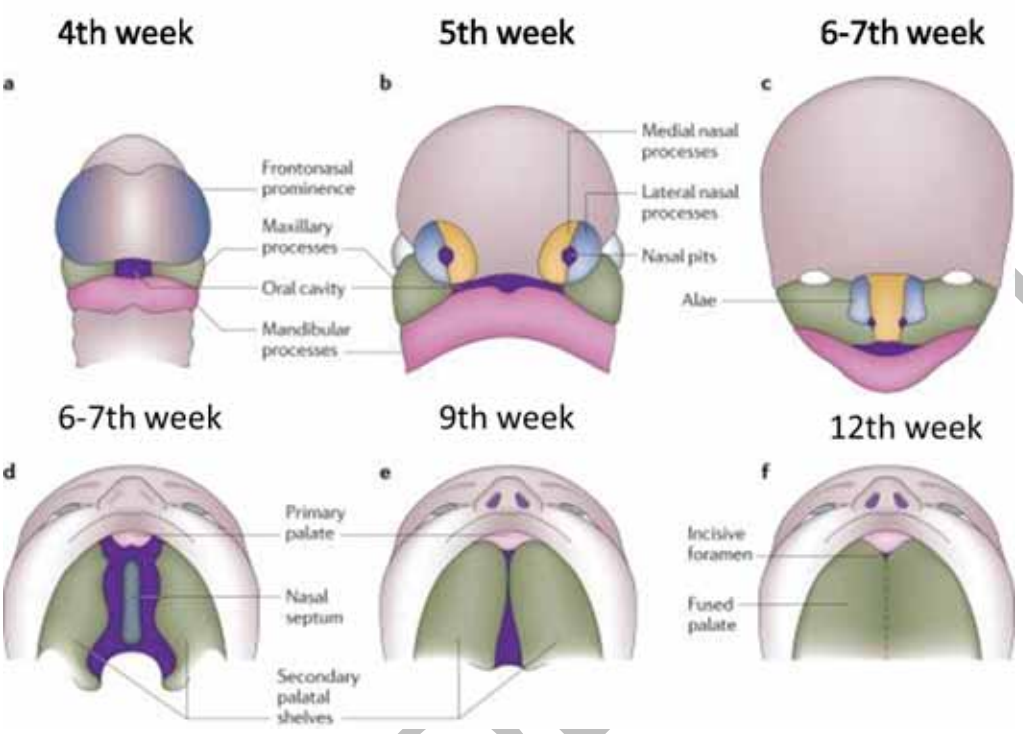


Figure 1. Development of the lip and palate.
a) By the 4th week, the primitive oral cavity is surrounded by the developing frontonasal, paired maxillary and mandibulary prominences. b) By the 5th week, invagination of the nasal placodes induces the formation of paired medial and lateral nasal prominences, delimiting the future nostrils. c) By the end of the 6th week, the medial nasal prominences and the maxillary prominences have fused to form the upper lip, the primary palate and the philtrum. Lateral nasal prominences form the nasal alae. Meanwhile, the paired mandibulary prominences merge to form the lower jaw. d) During the 6th week, palatal shelves grow from the maxillary prominences. They grow vertically on the side of the tongue. e) Paired palatal shelves elevate above the tongue and grow to contact each other and to begin fusion. f) By the 12th week, palatal shelves fused and separate the oral from the nasal cavity. adapted from [13].

B. Classification and epidemiology of cleft

1. Classification

a) Cleft types

Numerous classifications have been proposed. The Tessier classification is anatomic and purely descriptive, thereby avoiding to refer to etiopathogenic or developmental mechanisms [14]. It is commonly used by surgeons. Genetists rather use a classification that takes into account genetic, epidemiology and embryological studies of clefts, that classifies then into broad categories [15] (Figures 2 and 3).

I: Cleft lip (CL)

It affects the primary palate (the lip, the alveolar bone, the triangular intermaxillary palate) and extends posteriorly up to the incisive foramen -clinically to the incisive papilla-. According to the structures affected and the orientation, cleft lip can be divided into:

- Unilateral or bilateral: depending on the localization of the cleft relative to the philtrum: either on one side (left or right) or both sides
- Complete or incomplete: the cleft is complete when all three structures are affected
- Minimal involvement

II: Cleft palate (CP):

It affects the secondary palate and does not extend anteriorly to the incisive foramen. Cleft palate is divided into:

- Cleft of the hard or anterior palate: it affects only the ossified palate
- Cleft of the soft or posterior palate: it affects only the non-ossified palate. Bifid uvula and submucous cleft are subcategories of posterior cleft palate.
- Complete cleft palate: it affects both the hard and soft palate
- Minimal involvement

III: Cleft lip with or without cleft palate (CL/P):
This category comprises any combination of the first and second groups

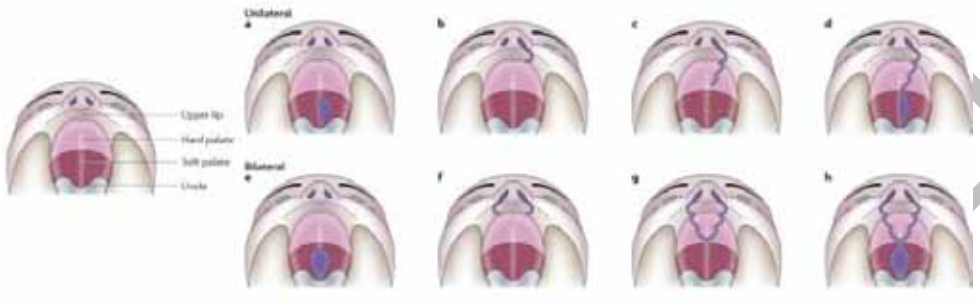


Figure 2: Types of cleft. Illustrative drawings of types of cleft lip and/or palate (CLP)
(A., B., C., D.) Unilateral cleft: (A.) Cleft of the soft palate; (B.) Cleft lip; (C.) Cleft lip and palate (hard); (D.) Complete cleft lip and palate (hard and soft); (E., F., G., H.) Bilateral cleft: (E.) Cleft of the soft palate; (F.) Cleft lip; (G.) Cleft lip and palate (hard); (H.) Complete cleft lip and palate. (Reprinted from [13]).

Based on epidemiological, embryological and recurrence risk studies, cleft lip (CL) and cleft lip with or without palate (CL/P) (Groups I and III) often aggregated, are considered to be different from cleft palate only (CP) (group II). They are traditionally studied separately. Occasionally however, both CL/P and CP can occur within the same family, suggesting at least some overlap in etiology of this two cleft categories. Such exceptions have been rarely reported in families with etiologic mutation in certain genes: i.e. *IRF6* [16], *MSX1* [17], *TBX22* [18], *FGFR1* [19], *TP63* [20]. When cleft palate occurs in CL/P patients, it is usually considered as a secondary effect of the cleft lip. Nevertheless, recent data suggest that CL and CL/P may have different developmental origins [21-23] and should, when feasible, be distinctly studied.

Minimal involvement
In addition to the overt cleft phenotype, "milder" degree of cleft malformation can be clinically noted. Minimal involvement of the lip is manifested by linear lips indentation(s) with or without nostril deformity -named "intrauterine healed cleft", "pseudocleft", "microforms"- or very rarely lateral fistulas of the upper lip.

Minimal involvement of the palate includes cleft uvula, submucous cleft palate (SMCP). Cleft uvula (bifid uvula) is described as uvula separated into two parts most easily seen at the tip. It is usually accompanied by SMCP, and been considered the marker for that condition. Submucous palatal cleft refers to a cleft of the hard palate with intact overlying mucosa, and, on inspection, a triad of (a) notching posterior border of the hard palate, (b) bifid uvula, and (c) muscular diastasis leading to a midline translucent zone or furrow in the soft palate. The mildest form is an occult submucous cleft palate, characterized by muscle malposition only [7].

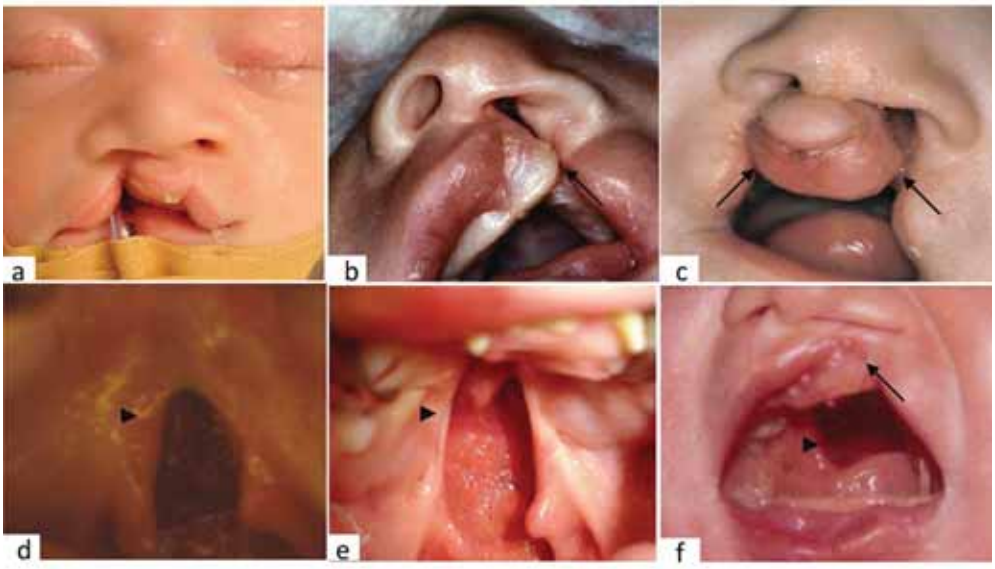


Figure 3: Cleft types.
(A.) Incomplete bilateral cleft lip. (B.) Unilateral cleft lip and alveolus (C.) Bilateral cleft lip. (D.) Cleft of the soft palate. (E.) Complete cleft palate (soft +hard). (F.) cleft lip and cleft palate (discontinuous cleft)
Pictures courtesy of Dr. B. Bayet, Cliniques universitaires Saint-Luc

b) Non syndromic versus syndromic clefts

Clefts can also be divided into "non-syndromic" (ns) (isolated) versus syndromic forms depending whether additional structural and/or developmental anomalies occur with the cleft. The majority of all cleft types, with approximately 70% of

CL/P cases and half of CP, are non-syndromic [24, 25]. The syndromic cases can be subdivided according to their etiology into chromosomal syndromes, single-gene or Mendelian disorders including over 500 Mendelian syndromes (see the OMIM web site), teratogenic and uncategorized syndromes. Non syndromic clefts can also be associated with minor anomalies (either microforms or subclinical physical features within the range of normal development) [23]. It is generally accepted that non syndromic clefts have no major dysmorphic signs and no more than two minor signs, such as found in 5% of the population.

The Pierre-Robin sequence

The well-recognized triad of micrognathia, cleft palate and glossoptossis is known as Pierre Robin sequence (PRS). Differences in definition and diagnostic criteria results in obvious difficulties in estimating its birth prevalence which ranges from 1/8,500 to 1/20,000, and is equal in both genders [26, 27]. These patients often suffer from respiratory and feeding difficulties that could be explained by the anatomical anomalies and/or the presence of a brainstem dysfunction [28]. A sequence is a cascade of anomalies that results from previous developmental anomalies or mechanical processes. The primary event that leads to PRS is still debated with micrognathia resulting in posteriorly placed tongue, preventing closure of the palatine shelves before the tenth week of gestation, and producing a U-shaped cleft palate, being the accepted dogma [29]. Others have considered that a primary defect in palatal closure would affect the maturation of muscles involved in deglutition and mastication and also would alter mandibular growth, resulting in PRS [30]. About one-quarter of infants with nsCP have Pierre-Robin sequence [23], and interestingly, with the exception of a shorter mandible in PRS, the facial morphology is similar to isolated CP suggesting that the mandibular retrognathia in PRS represents the end of a spectrum seen in isolated CP [7]. Some factors are suggestive of genetic basis in PRS. About half of PRS are isolated, and the two most common genetic syndromes associated to PRS are Stickler syndrome and velocardiofacial syndrome. When isolated, PRS is mostly sporadic and the etiology unknown. However, a family history of clefting and/or micrognathia has been reported in a significant subset of children with isolated PRS [31, 32] and 4 candidate loci and a few potential candidate genes for PRS have been proposed [32].

c) Phenotypic aspect of non-syndromic cleft

Phenotypic spectrum of non-syndromic cleft is more complex than previously thought. Finer subdivision of isolated clefts has recently been proposed based on the presence/absence of a variety of subclinical features (subphenotype) including minor structural variants [13], observed in either an individuals with cleft and/or their "unaffected relatives" [33]. Genetic studies also support the interpretation of mild forms of clefting as part of a spectrum of the cleft lip phenotype [34].

(1) Lip pits/prints

Whorl lip print patterns are circular grooves on the central upper lip and/or the left and right parts of the lower lip. Although they are typical symptom of the autosomal-dominant Van der Woude syndrome (OMIM 119300), their frequency is significantly elevated in non-syndromic CL/P individuals (14.8%) and their family members (13.2%), compared to unrelated controls (2.3%), and could be part of an expanded phenotypic spectrum [35].

(2) Dental anomalies

Dental anomalies, especially hypodontia, microdontia and supernumerary teeth, occur frequently in clefts [36]. The congenital absence of homolateral permanent incisor is the most commonly reported anomaly, but the missing teeth can also be outside the cleft area, particularly premolars and contralateral lateral incisor [37]. Dental anomalies are also more common in relatives of affected patients as compared to controls [38]. The facts that gene expression studies have shown that a number of genes co-localized in the developing teeth and palate [16, 39], that loci contributing to both cleft lip and dental anomalies have been identified [40], and that number of genes are associated with both clefting and tooth agenesis [41], support the hypothesis that dental anomalies contribute to the cleft phenotype. Therefore dental anomalies should be included in an overall assessment of the cleft phenotype [36].

(3) Orbicularis oris muscle defects

The orbicularis oris muscle (OOM) is the sphincter muscle that surrounds the mouth. It is deviated in CL/P patient as a result of the cleft lip, and either interrupted or narrowed in microform cleft [42]. OOM integrity can be visualized using ultrasound. It has been hypothesized that occult defects of the OOM may represent the mildest subclinical form of the lip portion of the CLP phenotype [43]. As demonstrated by using high resolution ultrasonography, the frequency of subclinical OO muscle discontinuities is higher in the non-cleft relatives (10.3%)

compared to the controls (5.8%) [43, 44], and is especially notable in males. Identification of mutations in the BMP4 gene in cases of subepithelial, microform, and overt cleft lip [34] supports the idea that subepithelial OOM defects are part of the spectrum of CL/P and ultrasound evaluation should be considered during clinical evaluation of families with clefts [34].

(4) Nasolabial fold discontinuity

The cleft phenotype (clinically identified cleft or occult orbicularis oris muscle defect) was hypothesized to affect not only the muscle itself but also the interaction with other facial muscles and the skin, and therefore influences movement in the midfacial region causing appearance-related changes. More nasolabial fold (NLF) discontinuity was found in CL/P patient and non-affected first degree relatives with OOM defect, compared to controls. There was no clear relation between NLF discontinuity and the side of the OOM defect, and the authors suggested that OOM and NLF discontinuity arise from the same genetic origin but have different effect on the soft tissue of the midface [45, 46].

(5) Face shape and three-dimensional facial image measurement

Subtle differences in the contour of the cranium morphology during embryogenesis have been known for many years [47]. In mice, the mechanisms by which the morphology translates into failed primary palate formation have been explored and morphological differences characterized in details. Since 1970 and the hypothesis of relation of face shape to susceptibility to congenital cleft lip [48], cephalometric and direct anthropometric studies have yielded variable and contradictory results [49].The introduction of non-invasive 3D surface imaging showed facial differences in non-affected parents of multiplex cleft families from that of the general population [50]. Increased nasal cavity width, a larger interorbital distance and midface retrusion have been identified in different studies as strong trait in non-affected relatives of patients with CL/P [34, 51]. Studies in twin discordant for nsCLP provide additional support that altered facial shape is a phenotypic marker for OFC susceptibility [52]. Hypothesis that genetic loci involved in nsCL/P also influence normal variation in facial morphology was supported by the association of one single nucleotide polymorphism (SNP) in 15q31 region with nose width, which could explain 2 % of nose width variation [53].

(6) Brain variants as assessed by MRI or by surrogate measures

Brain and face development are intimately related in both normal and pathological condition. Therefore it should not be surprising that abnormal brain development may be accompanied an abnormality in facial development. Neurological deficits, considered as mild, have been reported in otherwise healthy individual with isolated CL/P [54]. The hypothesis that these deficits could be a primary problem related to abnormal brain structure led to evaluate brain morphology in isolated CL/P compared to controls and MRI showed significant structural changes of the cerebrum and cerebellum in affected individuals [55-57]. As several candidate genes associated with CL/P (i.e *IRF6, MSX1, PVRL1*) are also associated with brain abnormalities and cognitive dysfunction, it reinforces the idea of genetic determinants involved both in nsCLP and formation of abnormal brain structure and function [58]. The occurrence of milder neuropsychological deficits and brain anomalies in "unaffected" relatives need to be determined.

(7) Speech pathology

A submucous cleft, like any defect in any aspect of nasopharyngeal anatomy (i.e. overt cleft palate) or physiology, may lead to velopharyngeal incompetence (VPI). It results in nasal escape of air and thus in a "nasal speech". Occult SMCP can be associated with functional disturbance leading to hypernasal speech. Clinical assessment of VPI can be very subjective; velar motion can be evaluated during sustained phonation of "a", and definitive diagnosis requires nasoendoscopy [33].

(8) Olfactory dysfunction

Several lines of evidence suggest that orofacial clefting is associated with deficits in olfactory ability. A significantly higher rate of olfactory dysfunction was observed in non-affected first-degree relatives of patients with nsCL/P in different studies [59, 60], providing evidence for the hypothesis that smell deficit is possibly a subclinical manifestation of the same underlying predisposition. Recently, relationship between facial characteristics and the smell capacity in non-affected first degree relatives of patients with nsCL/P was studied, showing correlation between olfactory dysfunction and a smaller upper nasal region [60].

2. Epidemiology

Orofacial cleft are among the most common birth defects with prevalence differences according to cleft type, ethnicity, gender, laterality and other factors. Although data differ somewhat from studies to studies depending for instance on type of source of information, methods of ascertainment and inclusion criteria, the average occurrence rate for CLP is around 1/700 live births.

a) Prevalence

(1) CL and CL/P

The average occurrence rate is said to be around 1/700. There is a marked racial predilection with the highest prevalence in Native Americans (3.6/1000 live births) and in Asians (Japan: 2.1/1000; China: 1.4/1000) and lowest in Blacks (0.3/1000). In the Caucasian population, the rate is intermediate: 1/1000, and the sex ratio is 2/1 (M/F) [61]. In general, the more severe the defect, the greater the male excess (ie: CL/P (2M/1F) is greater than CL (1,5M/1F), and bilateral is greater than unilateral) [7]. Most cases of CL are unilateral -with an amount of 90% if only CL is involved, 75% if CL is combined with CP-, and left-sided (70% if unilateral). Association with CP is seen in approximately 85% of cases of bilateral CL and 70 % of unilateral. CL can have various degree of completeness, involving the lip, the alveolar bone and the triangular shaped primary palate, and extension (i.e. extending into the nostril). CL can be associated with skin bridges (Simonart's band) in 10 to 30% [7].

(2) CP

CP does not vary significantly with geographical situation with an approximate frequency of 0.4/1000 births. There is a 2/1 female predilection for complete clefts of the hard and soft palate, ratio that approaches 1/1 for clefts of the soft palate only.
The incidence of cleft uvula is much higher than that of CP (1/2500 births). An incidence of 1/1200 to 1/2000 births has been reported for SMCP, with a 1/1 sex ratio [7].

Consistent associations between orofacial clefts and socio–economic status, or increased paternal or maternal age have not been established [23, 61].

b) Recurrence risk

The diagnosis of syndromic/non-syndromic CLP has significant implications for genetic counseling and recurrence risk estimation. The majority of syndromic clefts follows a Mendelian inheritance, and recurrence risk can be estimated accordingly. Nevertheless, because of varying penetrance in such syndromes, diagnosis can be difficult. In contrast, the majority of ns clefts (70%) are sporadic and have a complex etiology (with both genetic and environmental contributions), with an unknown mode of inheritance.
In families with CL/P, if one child is already affected, the risk of recurrence is 4%. In families with CP, the recurrence risk of having a cleft if one child is already affected is: 1.8%.
These recurrence risks increase with the number of affected individuals, and the degree of parental relationship [62] (Table 1).

Relative	CL (%)	CL/P (%)	CP (%)
Sibling	2.5	3.9	3.3
Half sibling	1.0	0.5	1.0
Parent	2.5	2.5	2.1
Offspring	3.5	4.1	4.2
Niece/nephew	0.9	0.8	1.1
Aunt/uncle	0.6	1.1	0.6
First cousin	0.3	0.5	0.4

Table 1. Frequency of oral clefts in relatives based on the proband's phenotype [62]

Family history and bilaterality of the proband's defect are significant predictors of recurrence risks among the siblings of CL/P probands (four-fold higher and two-fold higher respectively) whereas the sex of the proband and palatal involvement in the proband's defect are not [63].

c) Social aspect

Although not associated to a consequent rate of mortality in "developed" countries, CLPs represent a significant lifelong morbidity, requiring multidisciplinary treatment and care. Individuals with CLP can experience problems with feeding, speaking, hearing and social integration that can be corrected to varying degrees by

surgery, dental treatment, speech therapy, and psychosocial intervention [13]. CLPs also constitute an economic burden.

C. Etiology of clefts

Isolated CLPs were first recognized to have an inherited component by studies of familial recurrence risk of Fogh-Andersen [64]. Additionally monozygotic twins have a higher concordance rate (25 to 60%) compared to dizygotic twins (3 to 6%) [65]. Since the 100% concordance rate is not reached in monozygotic twins, it suggests the involvement of other factors, such as environment, to play a role in the occurrence of cleft. Some environmental risk factors have now been implicated in CLPs [13]. Moreover, few pedigree show clear-cut Mendelian inheritance and most cases appear to be sporadic. Consequently a model with a mixture of monogenic cases, probably dominantly inherited, and of multifactorial cases is favoured [66].

1. Genetics factors

In most population, review of segregation analyses of OFC concluded that genes of major effect, sometimes with evidence of multifactorial background effects, operate in the etiology of OFC [67]. Analyses based on evaluations of recurrence risk in OFC families in the aggregate estimated that from 2 to 15 genes of major effect are likely involved in OFC [68, 69]. To yield sufficient statistical power, statistical genetic analysis requires large numbers of families or large numbers of cases and controls. For years, research groups recruited large number of cases, helped by the facts that diagnosis is usually made at birth, and that 20% are familial cases.

a) Gene discovery

A variety of genetic approaches, often complementary, have been used to identify regions/genes for nsCLP: identification of chromosomal abnormalities or microdeletions in individual cases, direct sequencing, linkage analyses in families, and association studies.

(1) Syndromic genes

Numerous Mendelian syndromes have CLPs as major findings. Some of the genes responsible for these syndromes have also been implicated in non-syndromic

CLPs, such as *IRF6, FGF8/FGFR1, P63, TBX22, PVRL1* [25, 70], and can contribute through variable penetrance or action of different modifiers [69].

(2) Cytogenetic abnormalities

6% of malformations in liveborn children are associated with major chromosomal abnormalities (seen on karyotype); chromosomal anomalies are most frequently seen in syndromic OFCs [71] and numerous causal cytogenetic or submicroscopic genomic rearrangements have been found in OFC patients [7]. Analysis of breakpoints in patients with balanced rearrangements has identified several candidate genes in syndromic forms of OFC ie: *SATB2, FGFR1* and non-syndromic: *SUMO1, FAF1* [72-74].

SUMO1 was initially found to be disrupted in an isolated CL/P patient with balanced chromosomal translocation (46, XX, t(2;8)(q33.1;q24.3), leading to its haploinsufficiency [73]. Mouse model and expression studies confirmed its implication in lip formation, and later on association studies [75] and microdeletion [76] of *SUMO1* have been reported in CL/P patient.

By characterizing the breakpoints of a familial translocation, *FAF1* haploinsufficiency was implicated in a family with CP and PRS. Strong genetic association between nsCP and *FAF1* was reported in a multipopulation study and involvement of *Faf1* in palatogenesis was demonstrated by functional experiment in the zebrafish model [77].

Most CNVs (Copy Number Variation) have been found in syndromic forms of clefting; the most frequent interstitial deletion in human being del(22)(q11.2) is associated to a wide phenotypic spectrum including a palatal anomaly in ~50% of cases. Advances in techniques allow screening the genome more rapidly and with increased resolution. Recent studies have focused on the role of small CNVs in nsCLP and have implicated genes including: *BMP4, FGFR2, TFAP2A* and *SUMO1* [76, 78, 79]. Large cohort studies of patients with CNV, thanks to the development of publicly accessible databases of chromosome imbalances and phenotype in humans (i.e.: Decipher, see URL) are also of great interest in pointing relevant genes of OFCs [80].

(3) Sequencing

Another approach is to use targeted sequencing in cases with CL/P. Multiple "resequencing" studies of candidate genes have identified specific variants including coding mutations that are potentially etiologic in clefting candidate genes: *MSX1* [17], *BMP4* [34] , *FGFR1* and *FGF8* [81].

Of the new technologies, next generation sequencing is expected to make a huge impact on CLP discovery. Whole exome sequencing (WES) is currently applied under the significant peaks from genome–wide linkage and association studies [67]. Early successes have been seen in Mendelian traits, including syndromes that can include OFC: Kabuki syndrome and *MLL2* [82], Miller syndrome and *DHODH* [83] and Bartsocas-Papas syndrome and *RIPK4* [84]. Recently, WES studies conducted on a large family with Van der Woude syndrome (VWS) identified *GRHL3* as a second candidate gene for VWS [85].

(4) Animal models

Some candidate genes have been identified from animal studies, for example *MSX1* [86]. Animal models are of great help in the understanding of orofacial clefting conditions. Much of our knowledge on lip and primary palate development comes from chicks, and on the secondary palate from mouse models [4]. Spontaneous and genetically induced mutations and genetic manipulation offer numerous ways to study cleft in those models.

(5) Linkage and association studies

Linkage analysis always requires multiplex families. It tests for co-segregation of observed genetic markers and trait in a family, trying to determine whenever the marker and the familial trait are linked. Association studies tests for differences in frequency of markers in samples of affected or unaffected individuals from a population. The genome-wide resolution of SNP and CNV arrays have made possible more powerful approaches such as family-based linkage analyses as well as genome-wide association studies (GWAS).

Genome-wide linkage studies result for nsCLP in six regions -1q32, 2p, 3q27-28, 9q21, 14q21-24, 16q24- and fine-mapping of these regions showed significant results in 2 genes, previously associated in candidate gene studies : *IRF6* (1q32) and *FOXE1* (9q21) [67]. One single genome-wide linkage study of CP found suggestive evidence of linkage for three chromosomal loci (1p34, 2p24-25 and 12q21) [87].

OFC is one of the few birth-onset disorders to have been studied with the GWAS method. GWAS studies validated previously identified genes: *IRF6* and *FOXE1*, associated genes (*MAFB, PAX7, VAX1, ARHGAP29, NOG*) and loci (8q24, 16p13.3) for CL/P. The first GWAS for CP identified a common missense mutation in *GRHL3*. Some loci are population specific; some are not [88].

Interestingly, with the notable exception of *IRF6* and *FOXE1*, GWAS and genome-wide linkage results are different, emphasizing that the two approaches have different strengths.

Candidate genes/genomic loci	Technology and methodology
IRF6	Linkage, Association, Targeted sequencing, GWAS, Exome sequencing, Copy number variation
FOXE1	Linkage, Association, Targeted sequencing, GWAS
MSX1	Targeted sequencing
BMP4	Targeted sequencing
FGFR1	Targeted sequencing
FGFR2	Targeted sequencing
CRISPLD2	Linkage and Association
SUMO1	Mutation screen
TGFβ	Association
MAFB	GWAS
PAX7	GWAS
VAX1	GWAS
ARHGAP29	GWAS, Mutation Screen
Chr8q.24	GWAS
Chr16p13.3	GWAS
NOG	GWAS
GRHL3	Linkage, Exome sequencing
CDH1	Exome sequencing, Targeted sequencing
MGAM	Copy number variation
ADAM3A	Copy number variation

Table 2. A list of important candidate genes and loci identified for nsCLP

b) Main pathways implicated in lip and palate formation

Facial morphogenesis is a complex process that requires a spatiotemporal coordination of various cellular functions of both mesenchymal and epithelial cells. Disruption of any developmental steps can lead to an orofacial cleft. The identification of causal genes of OFC will help understand their underlying molecular mechanisms. Major signaling pathways have been pinpointed both from

human data and animal models like *FGF, WNT, TGF, BMP, PDGF* and *NOTCH* signaling. Many other genes have been implicated in craniofacial morphogenesis, as well as additional genetic or epigenetic mechanisms, and environmental factors.

(1) *FGF* signaling pathway

FGF (Fibroblast growth factor) signaling activation is regulated by their receptors and ligands, which have unique binding properties for individual FGF receptor (FGFR) isoforms. These FGFR are transmembrane tyrosine kinases activated upon ligand binding and receptor dimerization. Activation leads to target gene regulation through the canonical FRS/RAS/ERK pathway. FGFs regulate mesoderm establishment, neural crest patterning, myogenesis, and morphogenesis. FGF signaling occurs in both facial epithelial and mesenchymal cells, promoting cell proliferation and reciprocal epithelial–mesenchymal communication through region-specific expression of FGF receptor isoforms [89].

Many members of the FGF pathway contribute to CLP, and CLP are commonly seen in FGF-related craniosynostosis syndromes. Furthermore, the FGF family was associated to non-syndromic cleft and sequencing of 12 *FGF* family member found likely-disease causing mutations, with an estimated FGF signaling pathway contribution to 3-5 % of nsCLP [81]. Loss of function mutation of *FGFR1*, which causes Kallman syndrome, is also implicated in nsCLP [19, 81].

Murine inactivation of several Fgf (Fgf8, Fgf9, Fgf10 and Fgfr2) members generates CP only [90]. Fgf10 null mice exhibit wide open cleft palate but also, in some animals, ectopic fusion of the palatal shelves to the oral epithelium. Mesenchymal Fgf10 ligand, primary expressed in the anterior palate and acting through the epithelial Fgfr2b, regulates proliferation and apoptosis of palatal shelves epithelium and has a role in ensuring proper fusion, by maintaining minimal expression until the appropriate time [91]. FGF signaling participates in all different stages of palatogenesis and interacts with several other signaling induced by the expression of genes such as *SHH, WNT* and *TGF* [92].

(2) *WNT* signaling pathway

WNT ligands are secreted glycoproteins which, by binding to receptors in different combinations can activate at least three distinct signaling pathways: the β-catenin dependent (canonical) pathway, the β-catenin–independent planar cell polarity (PCP) pathway, and the β-catenin independent Ca^{2+} pathway. In the canonical pathway, Wnt ligands bind to their receptors, Frizzled and lipoprotein receptor-related proteins (LRP), to initiate an intracellular signaling cascade to the nucleus that is transduced by β-catenin. B-catenin accumulates and translocates into the

nucleus to bond lymphoid enhancer–binding factor/T cell-specific transcription factor (lef/Tcf) family member which will modulate targeted gene expression [93]. The Wnt signaling pathway regulates crucial aspect of embryonic development, and during craniofacial development plays critical roles in the generation, migration and proliferation of CNC cells in the facial processes.

Wnt9b is expressed in the ectoderm of the facial processes and in the epithelial seam between the fusing facial processes during lip fusion [94]. Wnt9b null mice shows reduced proliferation of mesenchymal cells caused by reduced FGF family gene expression and FGF signaling activity, which subsequently results in a failure of physical contact between the facial processes that leads to cleft lip and cleft palate [95]. Lrp6 is a key receptor of the canonical Wnt/β-catenin pathway and plays a major role in lip morphogenesis. Murine inactivation of Lrp6 induces bilateral CL, hypoplasia of the upper lip, cleft palate and mandibular defects. Its downstream effectors, including Wnt3a, Sostdc1, Radh3 and Msx1/Msx2 regulate both mesenchymal proliferation and epithelial fusion of lip primordia [96].

CP is present in several mutant mice. It has been reported that a number of components and modulators of Wnt canonical signaling, including Wnt2, Wnt3, Wnt4, Fzd6, Catnb, Dkk1, Lrp5/6, gsk3α/β, and Axin2 are extensively expressed in palatal epithelium, while the antagonists Sfrp2 and Sfrp4 are expressed in the mesenchyme [93]. B-catenin conditional mice show CP by persistence of MES. Tgfb3 is expressed in the palate MEE and β-catenin directly binds it. Downregulated expression and rescuing assays of Tgfb3 suggested that Wnt canonical signaling regulates palate fusion through Tgfb3 signaling transcription [93, 97, 98]. In addition, the Wnt canonical pathway signaling is involved in the initiation and patterning of palate rugae in close relationship with the SHH pathway. Loss of Wnt signaling in the palate epithelium blocks the rugae formation and altered Shh expression which in turn result in a unique anterior only cleft palate phenotype [98, 99]. Mice lacking Wnt5a or its non-canonical receptor Ror2 were found to exhibit CP due to defects in PS growth or elevation [100, 101]. Wnt5a expression is limited to the mesenchyme in a posterior-anterior gradient, and act as a chemoattractant causing cell to migrate from the posterior region to the anterior region of the palate [101]. Wnt11, another non canonical ligand only expressed in the MEE, has been suggested to mediate PSs fusion through a negative regulatory loop via Fgfr1 [102].

In humans, several Wnt ligands have been linked or associated with nsCLP [103-105]. Human Mutations in *WNT3* underlie autosomal–recessive tetra-amelia with cleft lip and palate [106]. Heterozygous mutation in *WNT5A* causes autosomal

dominant Robinow syndrome, and homozygous or compound heterozygous mutations in the *ROR2* cause the autosomal recessive Robinow syndrome.

(3) *SHH* signaling pathway

Sonic Hedgehogh (SHH) is the main actor of Hedghog pathway in craniofacial development. Binding of Shh to Patched (Ptch) transmembrane receptor derepresses Smoothened (Smo) and allows pathway activation, mediated downstream through Gli family transcription factors. Shh is expressed in the epithelium throughout palatogenesis. Expression is restricted to a striped pattern that corresponds to the rugae, which are suggested to act as centers that coordinate patterning within the palate limited to the nasal side of the palate epithelium [107]. No expression is seen in the medial epithelial seam during the fusion process [108]. A number of components within the Shh pathway are expressed in the epithelium and mesenchyme of the palatal shelves (PS), indicating the likely presence of both inter-epithelial and epithelial-mesenchymal signaling during palatogenesis. Shh coordinates a reciprocal epithelial-mesenchymal interaction which control the palate mesenchymal proliferation through: Fgf10/Fgfr2 signaling and Bmp/hand2 [109], during outgrowth of PS. Shh is also downstream of the Msx1 network that regulates cell proliferation in the anterior palate, and of canonical Wnt signaling in the developing palate.

Human mutations in several members of the SHH pathway -*SHH, GLI2, GLI3, PTCH, GAS1*- can lead to a variable degree of holoprosencephaly of variable severity, sometimes associated to CL/P [110]. Mutations in *GLI2* and *PTCH* are rare cause of isolated clefts [110, 111].

(4) *TGFβ/BMP* signaling pathway

TGFβ (Transforming Growth Factor β) superfamily includes TGFβs, the bone morphogenetic proteins (BMPs) which regulate embryonic patterning, the closely related growth and differentiation factors (GDFs) which regulate cartilage and skeletal development and the activins (Acts) and inhibins (Inhs). TGFβ superfamily ligands transduce their signal by bringing together type I and type II receptor serine/threonine kinases on the cell surface. This allows specific receptor II to phosphorylate the receptor I kinase domain, which then propagates the signal through phosphorylation of the transcriptional co-activators called Smad proteins (SMAD 2/3 upon TGBβ ligands and SMAD1/5/8 upon BMPs). These SMAD bind to the comediator, SMAD4 to form a complex which translocates into the nucleus to regulate target genes transcription, such as *MSX* [112]. Fine tuning of the

pathway is primarily achieved through the action of antagonists, co-receptors, and intracellular regulatory proteins [113].

In mice, Tgfβ1 and Tgfβ3 are expressed in the medial edge epithelial (MEE) of the PSs, whereas Tgfβ2 is restricted to CNC-derived mesenchyme underlying the MEE. Mice lacking Tgfβ3 exhibits failure of palatal fusion with persistent periderm cells covering the palatal shelves [114]. Tgfβ3 signaling acts through multiple intracellular pathways and transcription factors in the MEE and overlying periderm to induce cell cycle exit and subsequent degeneration of the MEE (Tgfβ3-Irf6-p63-p21), disrupt epithelial adhesion likely with loosening MEE and periderm cell adhesion (Snai1/2- E-Cadherin), and break down extracellular matrix (Mmp13) [115].

In Human, Loeys-Dietz Syndrome (LDS) -cardiovascular, neurocognitive, skeletal defects, and/or CP- is genetically heterogeneous with mutation in the *TGFBR1* gene (LDS2) , the *TGFBR2* gene (LDS2), the *SMAD3* gene (LDS3), associated with early-onset osteoarthritis), the *TGFB2* (LDS4) , the *TGFB3* gene (LDS5).

Irf6 mutant mice exhibit a hyper-proliferative epidermis that fails to undergo terminal differentiation, that lead to CP (resulting of occlusion of the oral cavity) and aberrant adhesion or fusion of PSs to mandible and/or tongue. Irf6 is a key determinant of the keratinocyte proliferation-differentiation switch, and has a key role in the formation of oral periderm [116, 117]. Irf6 is activated in the basal layer of the MEE cells, in addition to the periderm, prior to the palatal shelf fusion. It is a downstream target of Tgfβ3. Irf6 acts to downregulate p63 to induce p21-mediated cell cycle exit in the MEE [118] Irf6 expression contributes to the degeneration of the MEE [119] and to the induction of MEE apoptosis and dissolution of the MES [118, 120]. Tgfβ3 signaling and Irf6 function are needed for the activation of Mmp13 expression, that might contribute to break down the basement membrane of ECM and periderm desquamation [120]. Interestingly, an Irf6/p63 reciprocal regulation loop was also demonstrated in the palate, p63 being able to bind to an upstream enhancer element of Irf6 and activate its expression [121], while Irf6 induces the proteasomal-degradation of p63 [122].

In human, *IRF6* is a prime example of a gene that is mutated in a Mendelian disorder, van der Woude syndrome, but also associated with nsCLP [13]. P63 underlies several malformation syndromes that include CLP as a hallmark feature [123].

The BMP signaling pathway is implicated in a wide array of developmental processes, including cell proliferation, apoptosis and differentiation and tissue morphogenesis. Bmp ligands –Bmp2, Bmp4, Bmp7- and Bmp receptor-Bmpr1a-

are of particular interest in craniofacial development, Bmp2 and Bmp4- are expressed in a dynamic pattern in the facial primordia, with Bmp4 expression highly restricted to distal epithelia of the MNP, LNP, MNP and MxP. Bmp4 and Bmp2 expression pattern correlates with the expression of Msx1 and Msx2 in the underlying mesenchyme and furthermore their ectopic expression is able to induce expression of Msx1 and Msx2 in facial mesenchyme. These data suggest that Msx1 and Msx2 are downstream targets of the Bmp signaling. Conditional inactivation of Bmp4 in craniofacial primordia results in cleft lip only, whereas deficiency of Bmpr1a, a palatal mesenchyme receptor, results in CLP, showing that BMP signaling has distinct function in development of the lip versus the secondary palate [113]. Interestingly, specific deletion of Bmp4 in facial primordia leads, at E12, to mutant mouse embryos with a bilateral MNP and MxP, but only 22 % still displayed CL at E14, indicating some kind of in utero repair mechanism, maybe due to redundant roles of Bmp [124]. BMP signaling plays a role in the outgrowth of the anterior PSs. Specific inactivation of Bmpr1a in the palatal mesenchyme causes decreased proliferation resulting in partial CP and expression studies suggest a role of Msx1 and Fgf10 in the regulation of epithelial Shh, which also controls mesenchymal cellular proliferation [12]. Loss of function mutation in *Msx1* results in defects in palatal mesenchyme proliferation leading to cleft palate in mice [86].

In human, mutations in the *BMP4* have been implicated in non-syndromic patients with CL/P and subepithelial, microform and overt cleft lip [34, 125]. *MSX1* mutations have also been implicated in non-syndromic orofacial clefting with tooth agenesis [17, 126].

2. Environmental factors

Identification of environmental risks, as some can be easily manipulable, and particularly if the can be personalized with genetic covariates is an opportunity to provide better prevention [127].

First research on biochemical factors implicated in orofacial cleft goes back to 1950 [128], and numerous are now implicated [129]. Because identification of environmental components of clefting and studies of gene-environment interactions require large cohort studies, and although many environmental risk factors have been investigated [13, 61], relatively few strong associations are clearly established. Interestingly, pregnancy planning confers protection against non-syndromic OFCs. It reinforces the environmental risk factor notion [130].

(1) Maternal smoking

It has consistently been associated with increased risk of CLP, with an estimated overall odd ratio for having CLP of ~1.3 among offspring of mothers who smoked one to ten cigarettes/day. The risk seems to be stronger for CL/P (OR: 1.34) than for CPO (OR: 1.22), and increased with the number of cigarettes per day in CL/P [127]. Furthermore, study from the National Birth Defect Prevention on a large cohort corroborated the modest association between smoking and orofacial clefts, and identified specific phenotypes most strongly affected: bilateral CL/P and PR [131]. Assuming a causal relation between maternal smoking and orofacial clefts, 4% of all orofacial clefts and 12% of bilateral CL/P could be attributed to periconceptional maternal smoking (if their mother reported smoking at any time in the month before or the first 3 months of pregnancy).

Maternal passive smoking exposure, that is difficult to measure, was associated with a two-fold increase in risk of both CL/P and CP [132]. Thus, considering the number of women who smokes during pregnancy and who are exposed to environmental tobacco smoke, even the modest associations have a significant public impact.

Interestingly, a recent study conducted on mice showed that nicotine exposure caused improper palatal fusion, which resulted in persistent MES with no or partial disintegration of the seam. It might be due to impaired expression of different signaling pathways genes. It also suggests that nicotine does not induce complete cleft palate; however other constituents of tobacco cannot be ruled out for its induction [133]. Human CL/P and normal nicotine-treated fibroblasts had showed altered in vitro expressions of genes related to molecular signaling pathways and extracellular matrix metabolism [134].

(2) Alcohol

Exposure to maternal alcohol consumption has inconsistently been associated to isolated CLP. Several studies support an increased risk of CLP with “high” intake of alcohol (depending on the study: from four drinks/month to > five drinks/sitting), and with dose of alcohol, frequency and type [135-137]. Some authors have proposed an association restricted to CL/P while others have suggested a stronger relation with CP, or no association with either type [137]. Recently, deRoo and coauthors confirmed an elevated risk of oral cleft with five or more alcoholic drinks consumption per sitting during the first trimester of pregnancy (OR: 2.6), evident only in mothers or children who carried the *ADH1C* haplotype associated with reduced alcohol metabolism (OR: 3.0) [138]. The

teratogenic effect of alcohol may depend on the genetic capacity of the mother and fetus to metabolize alcohol.

(3) Diet

Based on observational and interventional studies, a role for maternal nutrition in orofacial clefts has been suggested. In most studies, maternal used of multivitamin supplements in periconceptional has been linked to decreased risk of orofacial cleft, and a meta-analysis associated multivitamin use with a 25 % reduction in birth prevalence of orofacial clefts [139]. However, composition of multivitamin complexes, associating multiple vitamins and minerals, are rarely reported, and responsibility (ies) of each component for this reduction is difficult to determine.

Folate deficiency causes cleft in mice [140]. Multiple studies have evaluated the role of folic acid in the occurrence and recurrence of orofacial clefts; while preventive effects of folic acid on orofacial clefts are commonly reported, the evidence remains generally inconsistent [141]. A recent large study on the effects and safety of periconceptional oral folate supplementation for preventing birth defects does not support any protective or negative effect of folic acid supplementation on orofacial cleft.

Other micronutrients have also been implicated in OFCs though the evidence remains weaker than that for folic acid. B1 and B6 deficiencies were associated with an increased risk of OFCs [142] as were myo-inositol and zinc [143], though data support roles for zinc deficiency in risk of oral clefts in populations in which zinc status is highly compromised [144]. Increasing preconceptional intake of nutrients predominantly present in fruits and vegetables reduces the risk of offspring affected by OFC [145].

(4) Caffeine

Caffeine ingestion during pregnancy has been suggested as a risk factor for cleft, and data are not consistent between studies. Coffee consumption was mildly associated with CL/P, but not with CP. When all sources of caffeine were considered (tea included) the evidence of an association between caffeine exposure and cleft was little. Data were not adjusted for confounding factor [146]. Others did not find any association with caffeine-beverage intake [147]. However, an association was observed association medications containing caffeine and isolated CL/P [147] (notably as many caffeine-containing medications contain other substances).

(5) Medication

Several medications have been linked to OFCs when taken during the first trimester.

Folate antagonists and certain other drugs that exhibit antifolate properties have been associated with an increased risk of CL/P and CP [148-150].

Positive associations with certain corticosteroids [151] during pregnancy have been reported but not confirmed in large cohort studies [152, 153]

(6) Maternal diseases

Some maternal diseases and obstetrics conditions have been associated to OFCs. Pre-pregnancy diabetes mellitus (PGDM) is a cause of embryofoetopathy, and a significant positive association was observed between PGDM and CL/P and CP as isolated or multiple defects [154]. Maternal obesity (Body mass index ≥ 30) was associated with a modest increased risk for both CL/P and CP in some studies; results that could be due to bias or residual confounding or that could represent a stronger underlying association [155]. Rarely, Amniotic bands syndrome, by disrupting fetal structures, have been causal for atypical OFCs [156]. Association of *IRF6* with clefts raises the possibility that viral infection in the first trimester of pregnancy might enhance risk of cleft [157].

(7) Environmental and occupational exposures

These exposures include hyperthermia [158, 159], stress, occupational exposures such as pesticide or chemical exposure of both parents [160, 161], ionizing radiation and infection [61]. Those exposures should continue to be assessed for possible role in clefting.

3. Gene-environment interactions

Gene variants can make an individual prone to a disease (susceptibility variant) or by decreasing his susceptibility or increasing his resistance against the disease less prone to (protective variant). A substantial number of studies suggest genetic and environmental factors may interact to modify risk to CLP. Certain metabolic pathways may have a role in the development of CLP.

Susceptibility to smoking exposure depends on biotransformation of toxic compounds by mother and embryo. Glutathione S-transferases GSTs are detoxification enzymes and markers in *GSTT1* and *GSTM1* genes appear to

influence risk of CL/P in the presence of mother smoking. The GSTT1 markers are gene deletion variants which confers an increased risk of CL/P if *GSTT1* null-genotype is carried by smoking mothers alone or together with the fetus compared to non-deleted non -smoking mothers (OR: 3.2 and 4.9 respectively) [162, 163], with stronger interaction effects in the CPO group (OR: 7,2). Moreover, combined absence of *GSTM1* and *GSTT1* in offspring of smoking mothers was associated with a higher risk of cleft (OR: 6.3). Interactions between smoking and several detoxification pathway genes : *CYP1A1, EPHX, NAT1* and *NAT2, NOS3* [163-165] have been reported Such data suggest that deficiencies in detoxification pathways may underlie some of the susceptibility to cleft [13]. Significant interaction was detected between maternal cigarette smoking, risk of cleft in offspring exposed to maternal cigarette smoking and variants in several developmental genes such as : *BCL3, BMP4, FGFR2, IRF6, MSX1, MYH9, SUMO1, RUNX2,TGFA, TGFB3* [75, 166-173].

With regards to maternal alcohol consumption, high dose of alcohol or "binge" drinking during the first trimester of pregnancy increases the rate of CLP in the offspring. Besides, the risk appears to be related particularly to mother or infant carrying specific haplotype in *ADH1C* (alcohol dehydrogenase 1C) encoding fast or slow alcohol-metabolizing phenotype. There was evidence of alcohol-related risk only in mothers or children who carried the *ADH1C* haplotype associated with reduced alcohol metabolism (OR= 3.0). Therefore the teratogenic effect of alcohol may depend on the genetic capacity of the mother and fetus to metabolize alcohol [137]. Risk estimates for high maternal alcohol consumption (≥4 drinks/month) were significantly elevated for CLP and were most elevated among infants with specific allelic variants in *MSX1* [170].

The 677 C --> T polymorphism in the 5-10 methylenetetrahydrofolate reductase (*MTHFR*) gene has been associated with non-syndromic CL/P in some populations. Mothers carrying the *MTHFR* 677TT genotype and who either did not use folic acid supplements periconceptionally or had a low dietary folate intake, or both, had an increased risk of delivering a CL/P child (odds ratio OR = 5.9; OR = 2.8, OR = 10.0, 95% respectively) [174]. Several studies have searched for interaction between vitamin supplementation and gene implicated in folate metabolizing genes and found inconstant results. Evidence suggesting an interaction between *IRF6* and maternal multivitamin use [169] and *ZNF533* and maternal vitamin use [175] as decreasing the susceptibility to isolated CL/P has also been reported.

II. AIMS OF THE STUDY

The general objective of this thesis was to further dissect the genetic etiopathogenesis of orofacial clefts by characterizing genetic variations implicated in non-syndromic clefts and Van der Woude syndrome. More specifically, by using whole exome sequencing we aimed to:

- Evaluate the implication of known syndromic CLP genes in non-syndromic forms.
- Study specific CLP phenotypes.
- Identify new cleft loci/genes.

III. MATERIALS AND METHODS

A. Patient selection

Samples for this study were drawn from a large cohort of 1300 nsCLP and nsCP index cases, of which 275 are multiplex families recruited over the years in collaboration with multidisciplinary teams.

Among familial cases, defined as familial history of CLP, that is closer than a 3rd degree affected relative, individuals with either a 22q11.2 deletion or an *IRF6* mutation were excluded. A total of 141 individuals from 77 non-syndromic CLP families, with pedigree structures consistent with autosomal-dominant inheritance and variable penetrance, and from 2 Van der Woude syndrome families were selected.

We also focused on a non-syndromic specific CLP phenotype named "discontinuous cleft", defined as an association of a cleft of the primary and the secondary palate with preservation of part of the intermaxillary palate and selected 8 individuals for study.

B. Whole exome sequencing

Different WES technologies were used with a first cohort of 89 individuals from 46 CLP families studied with SOLID sequencing technology, from Life Technologies, and a second cohort of 52 individuals from 33 CLP families studied using Illumina HiSeq 2000 technology. The 8 individuals with discontinuous CLP were studied using Illumina HiSeq 2000 technology.

As typically between 20,000 and 50,000 variants are identified per sequenced exome, a selection based on quality criteria is first made to avoid false-positive call of variants. Then in order to identify genes carrying rare deleterious variant, a first "variant priorization" is held on the frequency of the variants [unique or very rare in the general population- <1% in gnomAD, <0,5% in 1000G, in GoNL-] and on the effect evaluated by different predictive tools of the variant on the protein function [minimal effect variant, such as synonymous coding variant will be typically excluded]. In order to find the causative mutation among the 150-500 private variants left [176], additional strategies are applied: linkage, overlap, de novo, candidate, homozygosity and/or double-hit based strategy. The strategy choice mainly depends on the availability of well-phenotyped patients and family members, and on the mode of inheritance. They can be combined.

Following the flow-chart described below (Figure 4), we searched for rare potential pathogenic variants. The first strategy of prioritization used was a candidate gene list strategy. We considered a list of 89 genes (Table 3) implicated in both syndromic and non-syndromic CLP [25], nsCLP candidate genes [177] and a selection of CLP candidate genes that we gathered from publications [33, 80, 178, 179] including recently confirmed CLP genes, such as *ARHGAP29* [180] and *GRHL3* [85]. We secondarily searched for genes with candidate variants in ≥ 2 families.

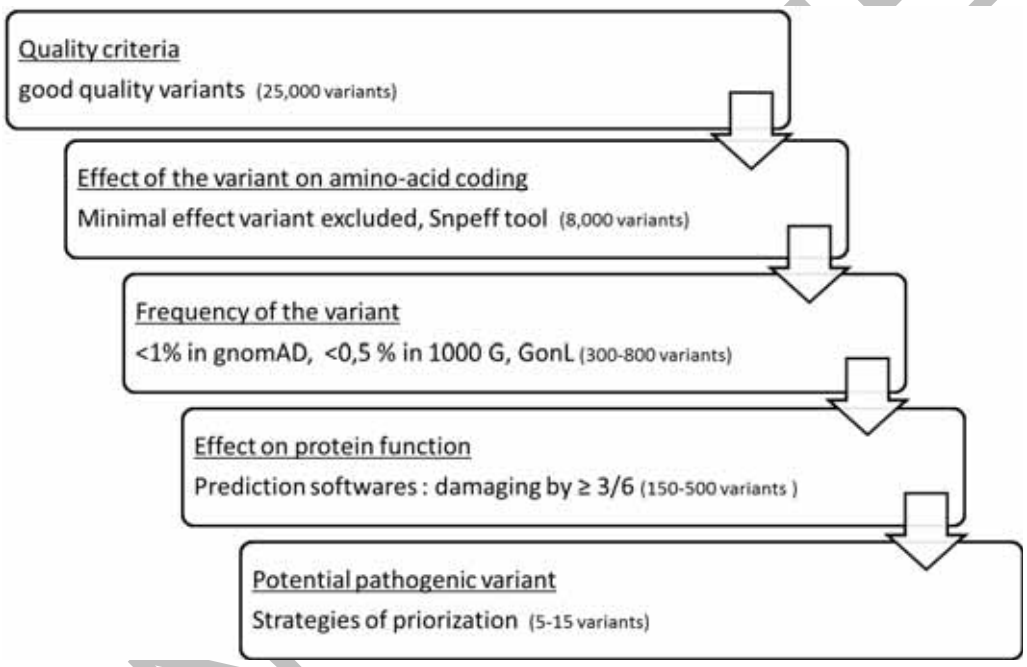


Figure 4. WES data analysis flow chart

ACTB	GLI2	RBM10
ARHGAP29	GLI3	RUNX1
ATG4C	GRHL3	RYK
BCOR	GRIM2	SATB2
BMP2	HYLS1	SF3B4
BMP4	IRF6	SHH
CDH1	JAG2	SIX3
CHD7	KLF4	SKI
CHRNG	LHX8	SLC26A2
COL2A1	MEIS2	SOX9
COL11A1	MID1	SPECC1L
COL11A2	MLL2	SPRY1
DLG1	MSX1	SPRY2
DLX4	MSX2	SUMO1
DHCR7	NIPBL	TBX1
DHCR24	NECTIN1	TBX22
EFNB1	NUDT6	TBX10
EFTUD2	OFD1	TCOF1
ESCO2	PAX9	TFAP2A
FAF1	PDGFC	TGFB3
FGF8	PHF8	TGFBR1
FGF10	POLR1C	TGFBR2
FGFR1	POLR1D	TGIF1
FGFR2	PQBP1	TP63
FLNA	PROK2	TULP4
FLNB	PROKR2	TWIST1
FOXC1	PTCH1	VAX1
FOXC2	PVR	WNT3
FOXE1	PVRL1	WHSC2
FOXF2	PVRL2	

Table 3. A list of 89 candidate genes for non-syndromic CLP

IV. RESULTS

A. Introductory remarks on publication I

"Whole Exome Sequencing unravels mutations in 10% of familial non-syndromic cleft lip and/or palate patients in genes mutated in well-known syndromes"

Mirta Basha, Bénédicte Demeer, Nicole Revencu, Raphael Helaers, Stephanie Theys, Sami Bou Saba, Odile Boute, Bernard Devauchelle, Geneviève François, Bénédicte Bayet, Miikka Vikkula

in Journal of Medical Genetics, Jul;55(7):449-458

The selection of patients for WES and WES data analysis started before I joined the laboratory. I helped MB with the WES data analysis and interpretation of the variants. Together with NR, I clinically reassessed patients and families, and collected DNA samples required for cosegregation studies.
This WES study on a large cohort of non-syndromic CLP families provides further evidence of contribution (~10 %) of syndromic CLP genes in non-syndromic forms. We suggest to use custom designed panels in diagnostic screen for families with nsCLP.

JMG Online First, published on March 2, 2018 as 10.1136/jmedgenet-2017-105110

ORIGINAL ARTICLE

Whole exome sequencing identifies mutations in 10% of patients with familial non-syndromic cleft lip and/or palate in genes mutated in well-known syndromes

Mirta Basha,[1] Bénédicte Demeer,[1,2,3] Nicole Revencu,[1,4] Raphael Helaers,[1] Stephanie Theys,[5] Sami Bou Saba,[6] Odile Boute,[7] Bernard Devauchelle,[8] Geneviève Francois,[9] Bénédicte Bayet,[10] Miikka Vikkula[1]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/jmedgenet-2017-105110).

For numbered affiliations see end of article.

Correspondence to
Professor Miikka Vikkula, Human Molecular Genetics, de Duve Institute, University of Louvain, Brussels 1348, Belgium; miikka.vikkula@uclouvain.be

Received 13 October 2017
Revised 2 February 2018
Accepted 12 February 2018

ABSTRACT

Background Oral clefts, that is, clefts of the lip and/or cleft palate (CL/P), are the most common craniofacial birth defects with an approximate incidence of ~1/700. To date, physicians stratify patients with oral clefts into either syndromic CL/P (syCL/P) or non-syndromic CL/P (nsCL/P) depending on whether the CL/P is associated with another anomaly or not. In general, patients with syCL/P follow Mendelian inheritance, while those with nsCL/P have a complex aetiology and, as such, do not adhere to Mendelian inheritance. Genome-wide association studies have identified approximately 30 risk loci for nsCL/P, which could explain a small fraction of heritability.

Methods To identify variants causing nsCL/P, we conducted whole exome sequencing on 84 individuals with nsCL/P, drawn from multiplex families (n=46).

Results We identified rare damaging variants in four genes known to be mutated in syCL/P: *TP63* (one family), *TBX1* (one family), *LRP6* (one family) and *GRHL3* (two families), and clinical reassessment confirmed the isolated nature of their CL/P.

Conclusion These data demonstrate that patients with CL/P without cardinal signs of a syndrome may still carry a mutation in a gene linked to syCL/P. Rare coding and non-coding variants in syCL/P genes could in part explain the controversial question of 'missing heritability' for nsCL/P. Therefore, gene panels designed for diagnostic testing of syCL/P should be used for patients with nsCL/P, especially when there is at least third-degree family history. This would allow a more precise management, follow-up and genetic counselling. Moreover, stratified cohorts would allow hunting for genetic modifiers.

INTRODUCTION

Cleft lip and/or cleft palate (CL/P) are the most prevalent craniofacial birth defects. They have an approximate incidence of 1/700 live births, varying with ethnicity, gender and cleft type.[1] Although CL/P is no longer associated with mortality in higher income countries,[2] it is a debilitating condition requiring expensive, invasive and often treatment until the adult age.

Syndromic CL/P (syCL/P) is typically caused by Mendelian mutation with a strongly deleterious effect on the function of a single protein, such as *IRF6*,[3] or by a deletion of a cluster of contiguous genes, such as del22q.11 (22qdel).[1] According to OMIM, there are 148 syndromes associated with CL/P.

Non-syndromic CL/P (nsCL/P), in contrast, is thought to have a complex aetiology, with multiple predisposing genetic variants acting in concert with intrauterine environmental effects. Epidemiological data suggest that if a child is born with a CL/P, the risk of re-occurrence in future siblings is calculated to be about 4%. The risk of recurrence is similar if one parent has CL/P, 17% if one parent and one child are affected and 9% when two children are affected.[4] Non-syndromic oral clefts constitute 70% of all CL/P subjects, of which 80% are sporadic and 20% are multiplex (familial) cases.[1]

A common hypothesis has been that a complex disease, such as nsCL/P, is due to accumulation of predisposing common variants with a minor allele frequency >5% in the affected subjects.[5] Association studies have been performed to identify them, either focusing on candidate genes or loci disrupted by karyotype aberrations in rare patients.[6 7] Genome-wide association studies (GWASs) have also been performed, and roughly 30 loci have been identified.[8–18] However, all these loci together explain only a small percentage of nsCL/P's heritability.[19] Thus, the enigma of 'missing heritability' remains and animates an intense debate.

We hypothesised that part of the missing heritability could be explained by rare (private) variants that are not assayed by GWASs. Such variants can be numerous in the population, with a medium-to-high penetrance. Moreover, there is high variability in expressivity in most single-gene oral cleft syndromes suggesting influence of genetic modifiers. Thus, rare variants in genes causing syCL/P could explain part of nsCL/P. We went out to look for such variants in a cohort of 84 individuals affected with nsCL/P, drawn from multiplex families (n=46).

METHODS

Patient selection

In close collaboration over 25 years with the multidisciplinary team at the Centre Labio-Palatin (Cliniques universitaires Saint-Luc), and a network of specialists practising in craniofacial centres worldwide, we have assembled a bio-bank of over 1300 samples of blood DNA from index subjects

To cite: Basha M, Demeer B, Revencu N, *et al*. *J Med Genet* Epub ahead of print: [please include Day Month Year]. doi:10.1136/jmedgenet-2017-105110

with different syndromic and non-syndromic CL/P. We have implemented systematic guidelines to collect samples uniformly. This included a detailed ascertainment of the clinical phenotype and family history by physicians, facilitated by a standardised questionnaire.

The database contains 275 families with CL/P, of which 43 families (16%) have an *IRF6* mutation, all diagnosed with a Van der Woude syndrome (VWS), either a priori (when lip pits were evident) or a posteriori (when an *IRF6* mutation was identified and minor atypical lower lip signs were identified in at least one family member following scrutinised clinical examination).[20] The remaining 234 families that were negative for *IRF6* mutation were candidates for whole exome sequencing (WES), in order to identify new causal gene(s). We selected a total of 84 individuals drawn from 46 families classified clinically as nsCL/P cases, for which we had enough high-quality DNA for WES. When possible, we chose two to four most distant relatives per family (26/46 families, 64 individuals). For 20 families, only the index patient was sequenced. The subphenotypes of these patients ranged from full-blown complete bilateral clefts of the lip and palate (BCLP) to a much more subtle velopharyngeal insufficiency (VPI). We had a total of 7 bilateral CLP, 16 unilateral CL/P, 8 unilateral cleft lip (CL), 33 cleft palate (CP), 7 CP-posterior, 2 submucous cleft palate (SMCP), 2 SMCP with bifid uvula and 9 VPI. All selected patients had standard karyotype and many had also been tested for 22qdel, and these tests were normal. All participants and their legal guardians (when necessary) gave informed consent for the participation in the study. This study was approved by La Commission d'Ethique Biomédicale Hospitalo-Facultaire, reference 2015/02NOV/572, of the medical faculty at University of Louvain, Brussels, Belgium.

Whole exome sequencing

Blood was drawn in EDTA tubes, and DNA was extracted from blood samples using Wizard Genomic DNA Purification Kit (Promega). WES was carried out by using 1 µg of genomic DNA per sample. Genomic DNA was fragmented into a library of small segments that can be uniformly and accurately sequenced in millions of parallel reactions. Exomes were captured using a commercial enrichment kit (Agilent SureSelectXT Human All Exon KitV5). Uncaptured DNA was washed off. Bounded DNA was clonally amplified and sequenced on Illumina HiSeq2000 to generate paired-end, 100 bp reads. Real-time image analysis and base calling were performed by sequence control software real-time analysis and CASAVA software V.1.8 (Illumina). We achieved a mean coverage of 61× without duplicates, with 81% of target covered at least 10×.

Data analysis

Generated reads were aligned to the human genome reference sequence (assembly GRCh37/hg19) by Burrows-Wheeler Aligner (see URLs). Picard Mark duplicates tool was used to flag duplicated reads (reads with a same start and end coordinates), which were subsequently removed from downstream analysis. SAM tools consolidated all BAM files (aligned reads). Local realignments around indels and base quality score recalibration were done with the Genome Analysis Toolkit (GATK) V.2.8. Variants were called using the GATK Unified Genotyper. Each called variant was annotated using Highlander software, an in-house bioinformatic framework (see URLs) (Helaers *et al*, under revision). On average, each patient had ±20 000–25 000 variants.

In silico analysis of called variants

We analysed variants in each family and in each sporadic subject independently. We retained variants for further analysis if they met the following criteria: (1) ≤1% in Exome Aggregation Consortium (ExAC), (2) ≤1% in 1000 genomes project, (3) ≤0.5% in Go-NL (genome of the Netherlands), (4) ≤1% in in-house controls (subjects affected with different condition than CL/P) and (5) estimated by visual inspection on Integrative Genomics Viewer/GATK scores as true variants.

As 'likely pathogenic' variants, we considered those with a high impact (canonical splice-site variants and those generating a premature termination codon, PTC (out-of-frame indels and nonsense variants)), or a moderate impact (missense), as predicted by SnpEff software. We prioritised variants in genes intolerant to loss-of-function (LoF) and/or variation, as estimated from the PLi values and z-scores derived from ExAC. We considered missense variants to have a deleterious effect when at least three out of six software (Sift, Polyphen, LRT, Mutation taster, Mutation assessor, FATHMM) predicted them as damaging. To maximise the conclusion that any novel, damaging variant would be causal, we focused on a list of about 500 (functional) candidate genes that we gathered from Jugessur and coworkers (ie, 357 biologically plausible, autosomal candidate genes for oral clefts),[21] augmented by novel genes presented at recent craniofacial conferences and or publications. We used the Genome Aggregation Database (gnomAD) to evaluate the frequency of the identified variants in a larger control population.

For all variants that satisfied our criteria, we performed co-segregation analysis on all the affected and non-affected family members when available. Approximately 15–20 variants/family and ~20–30 variants/sporadic subject were PCR-amplified followed by sequencing on an ABI 3130XL genetic analyser (Applied BioSystems).

To evaluate genotype–phenotype correlations in the identified genes, we gathered all intragenic published mutations from HGMD, denovo-db and current literature. We considered only missense, nonsense, canonical splice sites and exonic small indels. These variants were dichotomised into either missense/in-frame indels or LoF (out-of-frame indels, nonsense and canonical splice-site variants).

Validation of mRNA stability

RNA was extracted from blood samples (lymphocytes) with TriPure (Roche) and retrotranscribed using RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas) with random hexamers. PCR-amplification was done on cDNA using a primer pair, forward in exon 2_3 (CCACCTGGACGTATTCCACT) and reverse in exon 5 (ACTTGCCCATCTCTGGTTTC). Amplicon's size (433 bp) was evaluated with an agarose gel, using GeneRuler 100 bp (Thermo Scientific) DNA ladder.

RESULTS

Family 1: *TP63* mutation

In family CLP-1055, we identified a 2 bp insertion (NM_003722.4: c.819_820dupCC) in exon 6 of *TP63* in the father and the proband (figure 1—F1). This mutation leads to a frame-shift with a PTC (NP_003713.3: p.GIn274fs*4:). The insertion is not known in the gnomAD containing 246,134 *TP63* alleles. The gene is intolerant for LoF mutations as evidenced by the PLi value of 0.98 (divergence between observed and expected counts for LoF changes) provided by ExAC. In all the 12 known *TP63* coding transcripts, the frame-shift is predicted to induce PTC in an evolutionary conserved DNA binding

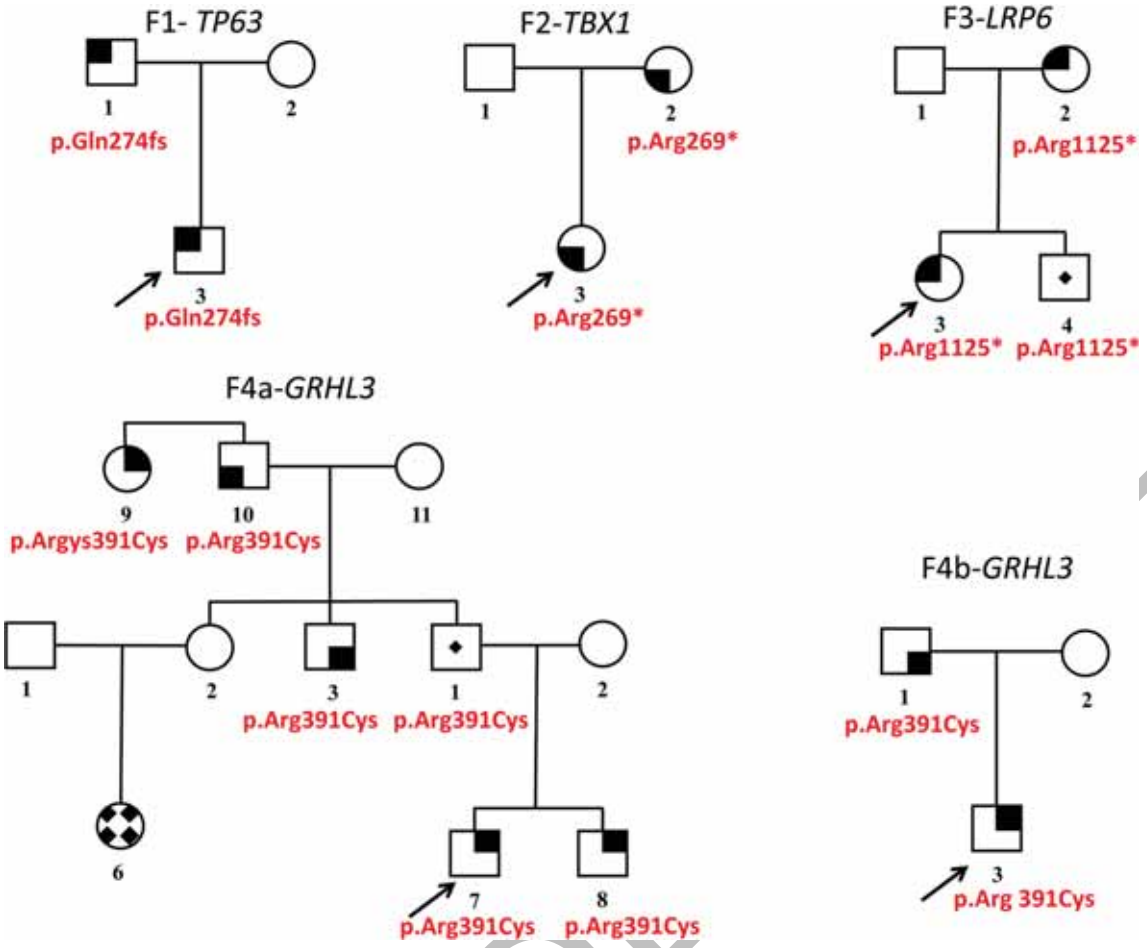


Figure 1 Five pedigrees with four mutations in cleft lip and/or palate (CL/P) syndrome causing genes. Family number and mutated gene marked above the pedigree (F-#- gene). Mutation is given underneath each carrier/affected. Gender symbols shaded with solid black areas indicate the subphenotype of CL/P. Upper right, CL/P; upper left, cleft palate (CP) or Pierre-Robin sequence; lower left, submucous cleft palate±bifid uvula; lower right, velopharyngeal insufficiency and/or hypernasal speech. Gender symbols with a central dot indicate carriers. A female symbol with a solid diamond pattern indicates a phenocopy CP. An arrow marks proband in each family.

domain (figure 2). Further, mRNA studies demonstrated that the mutant allele underwent nonsense-mediated-mRNA decay (NMD) as only the wild-type allele observed in mRNA obtained from the patient's lymphocytes and lymphoblasts.

The proband had a unilateral, right-sided CLP. He had no limb anomaly, no ectodermal dysplasia and no cardiac malformation. Follow-up until the age of 3.5 years showed growth and development within normal limits. His father had a unilateral, left-sided CLP. Both were re-examined for minor symptoms of *TP63* disorders, yet they sustained their initial classification as nsCL/P.

Family 2: *TBX1* mutation

In family CLP-678, we identified a nucleotide change (NM_080647.1: c.805C>T:) in exon 6 of *TBX1* in the mother and the proband (figure 1—F2). This nonsense mutation (NP_542378.1: p.Arg269*: NP_542378.1) is located in the T-Box domain (figure 3). The substitution is unknown in gnomAD containing 277 208 *TBX1* alleles. As *TBX1* mRNA is not expressed in lymphocytes or lymphoblasts, mRNA stability could not be studied.

The index subject was born at 35.5 weeks of gestation. She had neonatal regurgitations and bradycardia. A diagnosis of VPI with severe hypotonia of the velopharyngeal musculature was made, but the 22q11 deletion test was normal. Follow-ups showed normal development, no particular facial gestalt, yet with a persistent hypernasal speech (a consequence of VPI), despite receiving speech therapy. Apart from a neonatal patent ductus arteriosus, the index case had no cardiac structural defect.

The proband's mother also had isolated VPI. She underwent surgery at the age of 23 years, with partial improvement. Cardiac ultrasound did not reveal any cardiac defect.

Family 3: *LRP6* mutation

In family CLP-451, we identified a nucleotide change (NM_002336.2: c.3373C>T:) in exon 15 of the LDL receptor related protein 6 (*LRP6*) in three individuals (figure 1—F3). The substitution is not known in gnomAD containing 277 132 *LRP6* alleles. This nonsense variant (NP_002327.2: p.Arg1125*:) is located in a conserved YWTD repeat in the fourth β-propeller domain (figure 4).

The proband had a BCLP with missing upper lateral incisors, and her mother had a bilateral CL. Proband's brother is an unaffected carrier of the substitution. We could not exclude oligodontia (>6 missing teeth) in any of the three mutation carriers since the family was not available for reassessment.

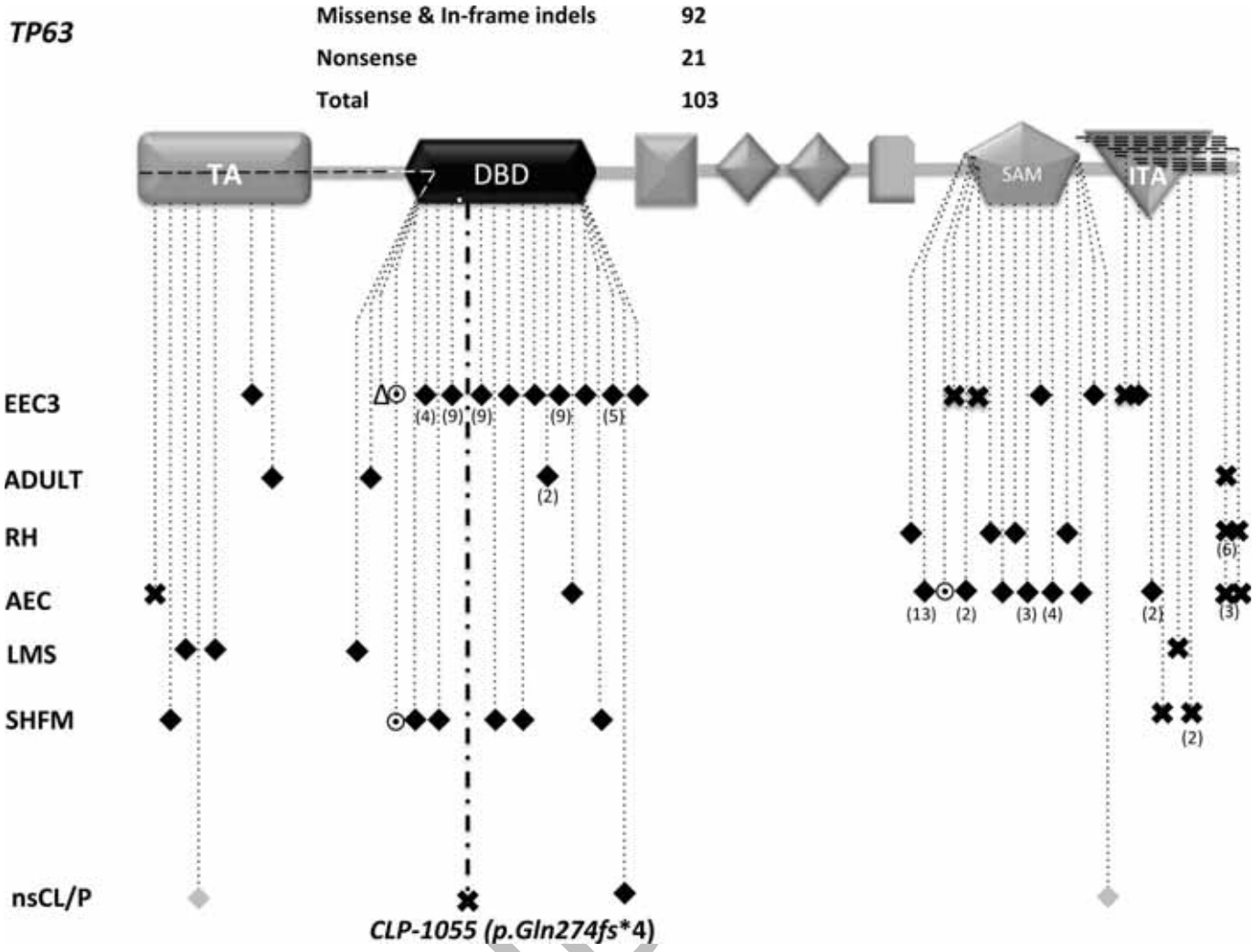


Figure 2 *TP63* protein structure; domains/motifs represented by different geometrical shapes (all scaled). ADULT, acro-dermato-ungual-lacrimal-tooth syndrome; AEC, ankyloblepharon–ectodermal defects–cleft lip/palate; DBD, DNA-binding domain; EEC3, ectrodactyly, ectodermal dysplasia and cleft lip/palate syndrome 3; ITA, inhibitor of transactivation domain; LMS, limb–mammary syndrome; RH, Rapp-Hodgkin syndrome; SAM, sterile alpha motif; SHFM, split/hand–foot malformation; TA, transactivation domain. Symbols: ✖, loss-of-function variants (exonic premature termination codon, out-of-frame indel and canonical splice-site variant); dashed line across protein, out-of-frame indel from start to end; ◉, in-frame indel; ◆, missense variants (neomutation, or present in a minimum of two affected, or functionally validated); ◆, other rare missense variant; dashed line across protein ending with Δ, a large 5′ deletion of several exons; number in parentheses, number of occurrences in index patients, mutations in close proximity grouped (see online supplementary table S1).

Families 4 and 5: *GRHL3* mutation

We identified an identical nucleotide change (NM_198174.2: c.1171C>T:) in exon 9 of Grainyhead-like transcription factor 3 (*GRHL3*) in two families CLP-398 (figure 1—F4a and CLP-986 (figure 1—F4b). The substitution is listed once in gnomAD containing 244 128 *GRHL3* alleles. This variant leads to an arginine-to-cysteine substitution (NP_002327.2: p.Arg391Cys) in an evolutionarily conserved DNA-binding domain (figure 5). This change has been reported as a de novo variant in two patients, one with a CLP and lip pits,[22] the other with spina bifida.[23] In vitro and in vivo studies demonstrated a dominant negative effect.[22 24]

The proband in family CLP-398 had a Pierre-Robin sequence (PRS) (figure 1—F4a). His brother had a milder PRS. On the paternal side, an uncle had a high-arched palate with a SMCP and missing upper lateral incisors. A paternal great-aunt had a surgically repaired CP. The grandfather has hypernasal speech, according to family history. The father of the proband is an unaffected carrier.

The proband in family CLP-986 had a posterior CP (affecting the soft palate). His father had a SMCP with a bifid uvula (figure 1—F4b). Abnormalities of the lower lip, such as lip pits associated with the VWS, were excluded following detailed clinical examination in both families.

DISCUSSION

We identified altogether 5 out of 46 index patients screened to have a clear mutation in a gene mutated in specific syndromes associated with oral clefts. The pick-up frequency among them is thus about 10% (5/46). These were all subjects with family history of nsCL/P. In total, there were 15 mutation carriers. Penetrance was high 13/15 (87%). Apart from family CLP-451, all patients were clinically re-examined for minor signs of CL/P syndromes. The assessment confirmed the isolated occurrence of

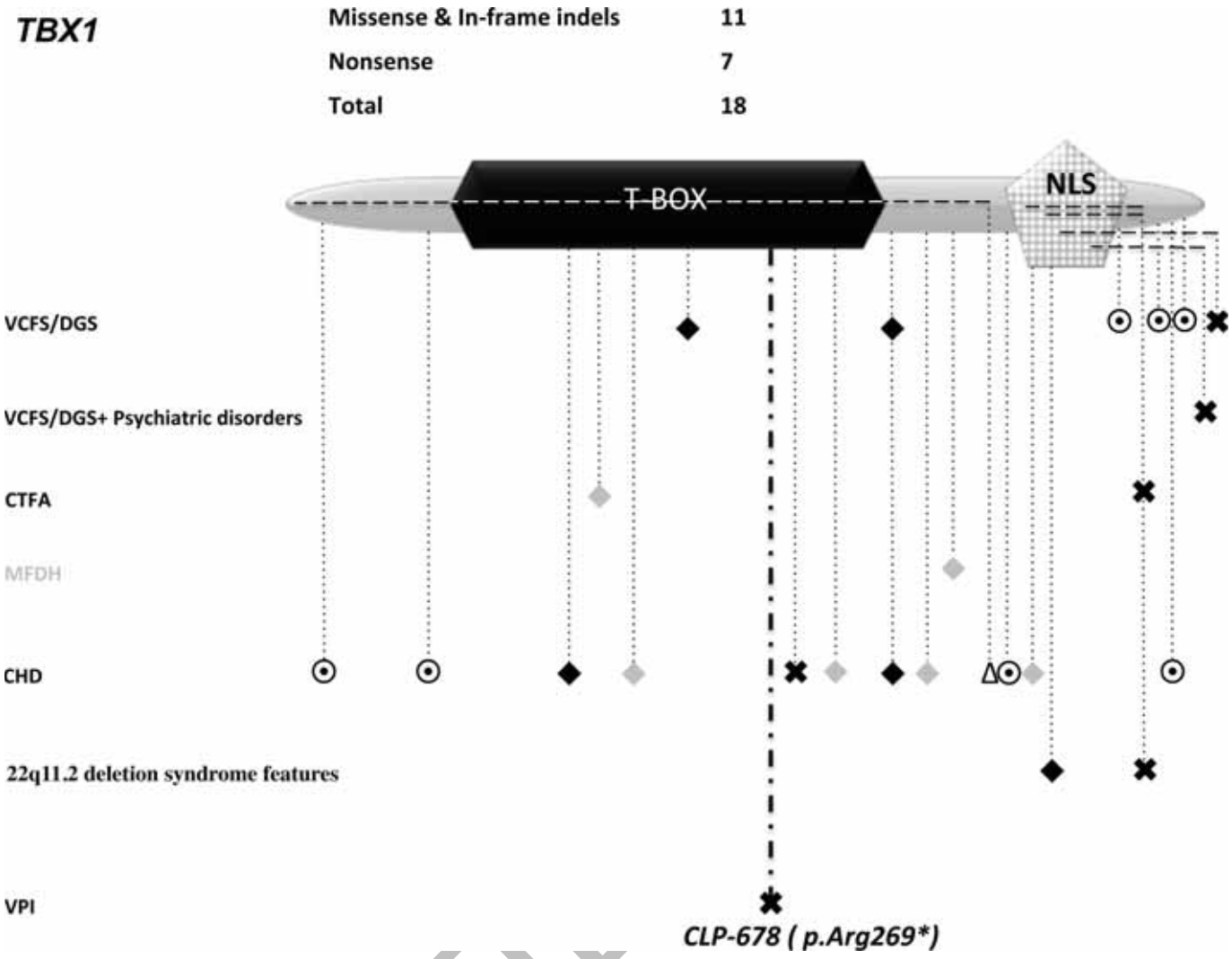


Figure 3 *TBX1* protein structure; domains/motifs represented by different geometrical shapes (all scaled). CHD, congenital heart defect; CTFA, conotruncal face anomaly syndrome; MFDH, midline facial defects with hypertelorism; NLS, nuclear localisation signal; T-BOX, DNA-binding domain; VCFS/DGS, velocardiofacial syndrome/DiGeroge syndrome; VCFS/DGS + psychiatric disorders such as depression and Asperger syndrome; VPI, velopharyngeal insufficiency. Symbols: ✖, loss-of-function variants (exonic premature termination codon, out-of-frame indel and canonical splice-site variant); dashed line across protein, out-of-frame indel from start to end; ◉, in-frame indel; ◆, missense variant (neomutation, or present in a minimum of two affected, or functionally validated); ◆, other rare missense variant; dashed line across protein ending with Δ, a large 5′ deletion of several exons (see online supplementary table S2).

CL/P. This underscores the extreme variability in expressivity in phenotypes caused by mutations in these genes, which can range from a full-spectrum syndrome to a most subtle CL/P subphenotype such as VPI.

In total, four mutations in four different genes were identified in five families (one mutation was shared by two unrelated families). The four mutations co-segregated with the oral cleft in an autosomal dominant manner in all families. They belong to the less frequent group, when known mutations are dichotomised into LoF (PTC or variant in a canonical splice site) and missense mutations. In *TP63*, *TBX1* and *LRP6*, in which we identified a LoF mutation, LoF variants account for only 21/103 (figure 2), 7/18 (figure 3) and 9/28 (figure 4), respectively. In contrast, in *GRHL3* (figure 5), in which we identified a missense mutation, validated missense variants[23 24] account for only 7/20.

In two of the four genes, localisation of our mutations did not follow the typical grouping of mutations by type. In both *TP63* and *TBX1*, the few known LoF mutations cluster to the 3′ end of the gene (figures 2–3 and online supplementary tables S1 and S2) and have been identified in patients with a full-blown syndrome, such as EEC3 (ectrodactyly, ectodermal dysplasia MIM #604292, and cleft lip/palate syndrome 3), ADULT (acro-dermato-ungual-lacrimal-tooth syndrome MIM #103285), RH (Rapp-Hodgkin syndrome MIM #19400), AEC (ankyloblepharon–ectodermal defects–cleft lip/palate MIM#19400), LMS (limb–mammary syndrome MIM #603543), SHFM (split/hand–foot malformation 4 MIM # 605289) or (VCFS/DGS), respectively (figures 2–3). Our LoF mutations localised to the DNA-binding domain, in the middle of the two proteins. Thus, there seems to be a genotype–phenotype correlation regarding mutation type and/or localisation for these genes.

TP63 encodes a transcription factor, which is a master regulator of epithelial lineage commitment during and after development. Heterozygous germ-line mutations in *TP63* underlie at least six disorders with overlapping phenotypes, which share ectrodactyly, ectodermal dysplasia, cleft lip/palate (EEC), lacrimal duct obstruction, hypopigmentation and hyperplastic breasts and/or nipples to varying degrees. Mutations are dispersed in all

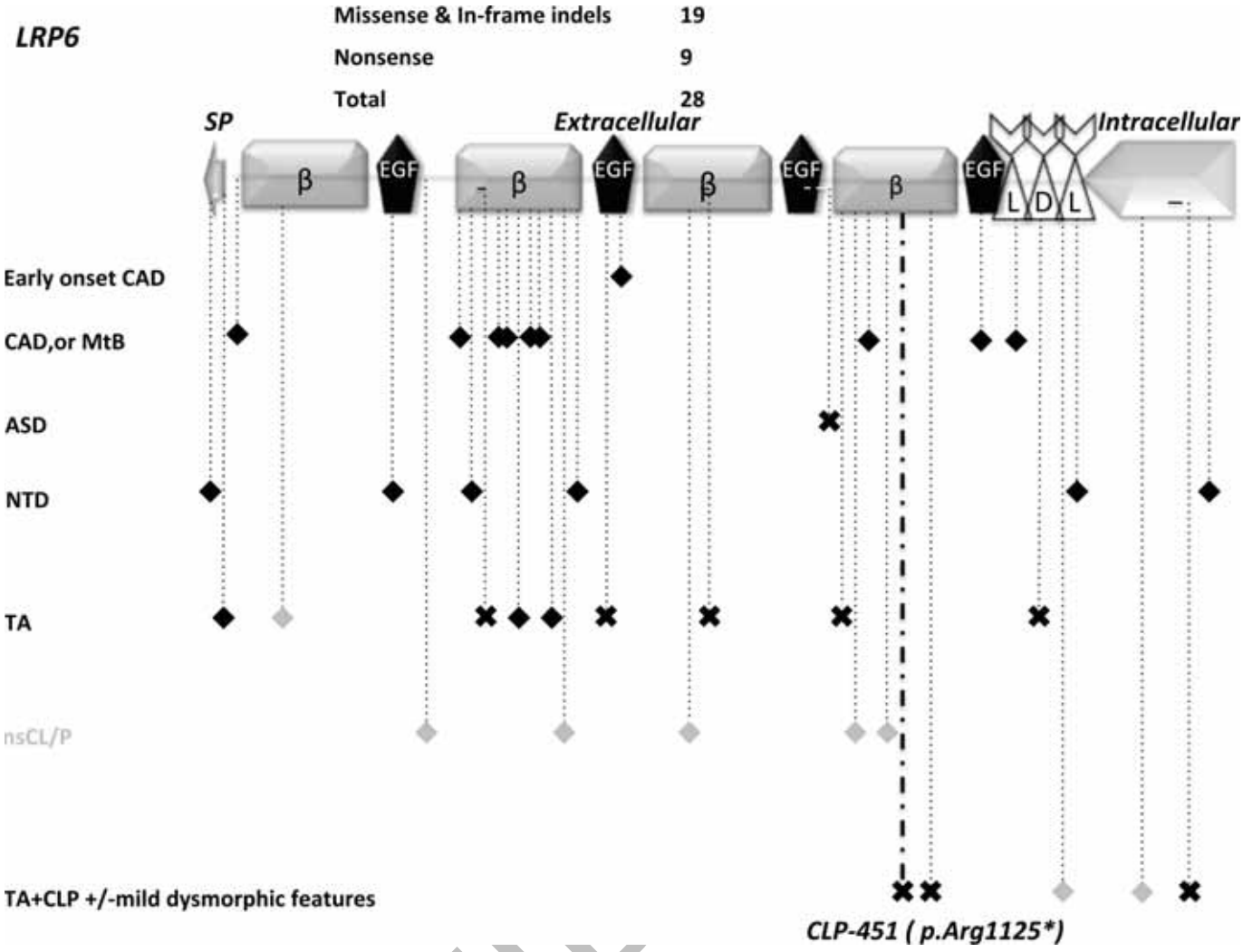


Figure 4 *LRP6* protein structure; domains/motifs represented by different geometrical shapes (all scaled). ASD, autism spectrum disorder; CAD, coronary artery disease; CLP, cleft lip and palate; MtB, metabolic syndrome; nsCL/P, non-syndromic cleft lip and/or palate; NTD, neural tube defect; TA, tooth agenesis; TA+CL/P, combined tooth agenesis with CLP and mild dysmorphic features. Extracellular part: β, β propeller repeats; EGF, epidermal growth factor-like domain; LDL, LDL type A-like receptor; SP, signal peptide. Symbols: ✖, loss-of-function variants (exonic premature termination codon, out-of-frame indel and canonical splice-site variant); dashed line across protein, out-of-frame indel from start to end; ◆, missense variant (neomutation, or present in a minimum of two affected, or functionally validated); , other rare missense variant (see online supplementary table S3).

functional domains of *TP63*, with a well-established genotype–phenotype correlation (figure 2). Penetrance is incomplete and variable expressivity is observed. EEC point mutations cluster in the DNA binding domain (DBD) and have a disruptive effect on DNA binding characteristics of all expressed isoforms. The effects of these mutations on the counterpart wild-type allele are thought to be dominant negative.[25] To our best knowledge, this is a first reported indel mutation within the DBD domain. Yet, the position is not that important, as the transcribed allele undergoes mRNA decay. A larger deletion extending from exon 1 to exon 4 (encompassing most of DBD) was reported in a patient with EEC3.[26] It is not known if a stable *TP63* lacking most of the DBD domain is produced starting from one of downstream in-frame ATGs and what kind of activity such a short form would have.

Whether *TP63*'s haploinsufficiency could also be the pathomechanism for some types of *TP63* disorders has been a subject of debate and was originally refuted based on inferences from mouse model phenotypes when compared with humans. Tp63$^{-/-}$ mice recapitulate the phenotypic spectrum observed in humans with a *TP63*$^{+/-}$ mutation,[27] while tp63$^{-/+}$ mice nor patients with a 3q27del do not exhibit ectodermal defects.[27] However, recent publications report patients with ectodermal dysplasia and high-arched palate,[28] among other symptoms with a 3q27.3del and a variable expressivity in a CP family with 3q28del.[29] This suggests that *TP63* haploinsufficiency can be a driving mechanism behind clefting. Therefore, the identified *TP63* LoF allele, which manifests as isolated nsCLP, suggests it to be a haploinsufficient allele.

TBX1 maps to 22q11.2, a region frequently deleted in humans (~1/4000 live births). It is associated with DiGeorge syndrome (DGS)/velocardiofacial syndrome (VCFS) (MIM 192430/188400). It regulates a number of genes via epigenetic modifications.[30 31] *TBX1* mutations manifest with pharyngeal arch-related developmental defects, affecting the heart, thymus and parathyroid gland. *TBX1* mutations cause a recognisable spectrum of craniofacial anomalies such as CP, or VPI with typical facies and a predilection to develop neuropsychiatric disorders.

Phenotypic discordance was observed in cases with an identical deletion and also among monozygotic twins.[32] Deleterious variants located within the Brachyury T domain (a well-conserved

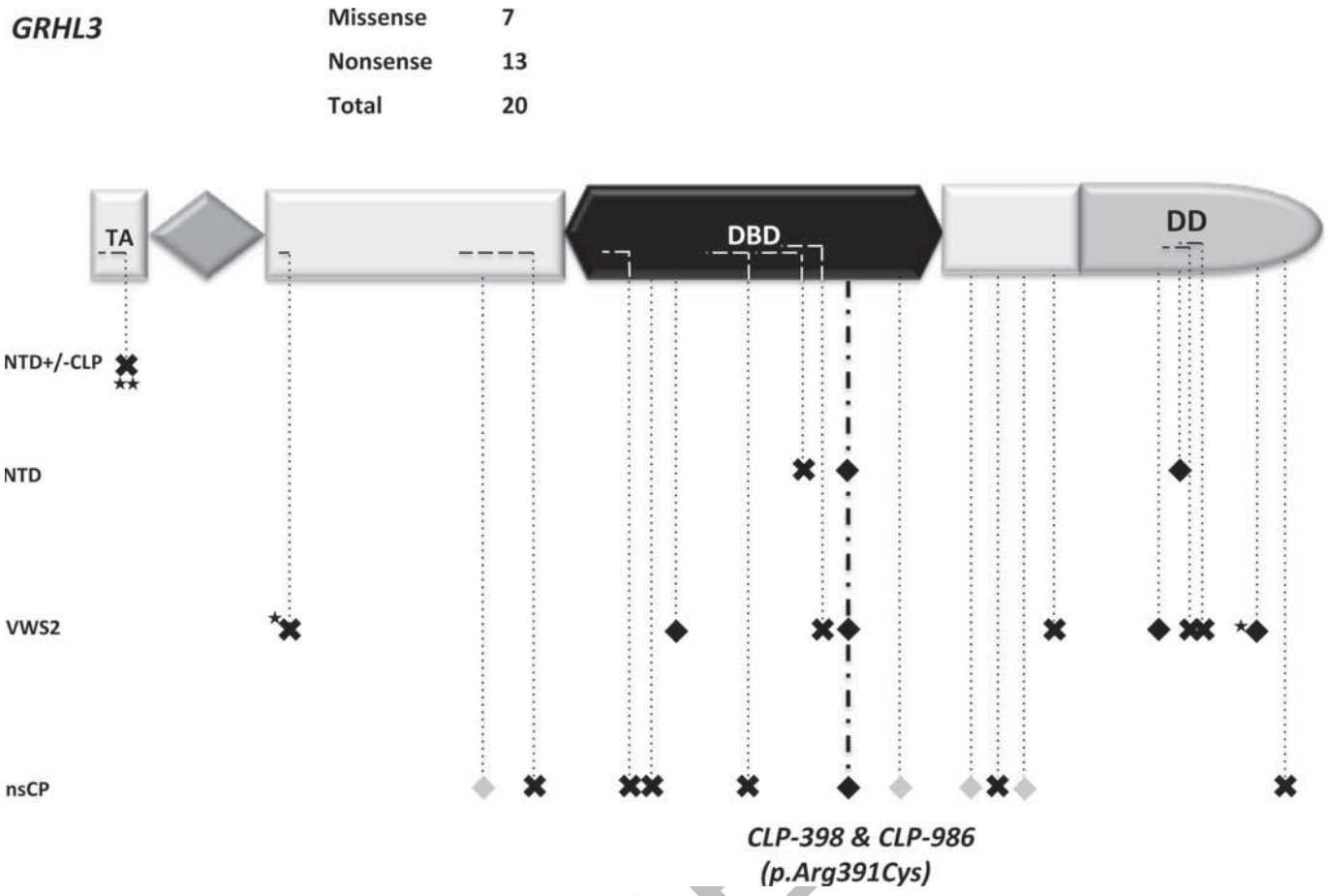


Figure 5 *GRHL3* protein structure; domains/motifs represented by different geometrical shapes (all scaled). CLP, cleft lip and palate; DBD, DNA-binding domain; DD, dimerisation domain; NTD, neural tube defect; nsCP, non-syndromic cleft palate; TA, transactivation domain; VWS2, Van der Woude syndrome 2. Symbols: ✖, loss-of-function variants (exonic premature termination codon, out-of-frame indel and canonical splice-site variant); dashed line across protein, out-of-frame indel from start to end; ◆, missense variant (neomutation, or present in a minimum of two affected, or functionally validated); ◆, other rare missense variant; ★, compound heterozygous mutations, identified in a single family; ★★, homozygous mutation (see online supplementary table S4).

DBD, shared with T-Box family members) are associated with DGS/VCFS (figure 3 and online supplementary table S2) and some variants are associated with isolated congenital heart defects (CHDs) (figure 3 and online supplementary table S2). In vitro investigations by luciferase assay demonstrated that the only identified PTC in the DNA-binding domain of *TBX1* p.Glu277*, which co-segregated in a family with CHD, is a LoF variant.[33] This variant is distal to our p.Arg269* variant, which manifests as isolated VPI without any cardiac defect. There are another four reported frame-shift mutations towards the 3′ end of *TBX1*, all disrupting the nuclear localisation signal (figure 3). These mutants, if stable, could inhibit *TBX1* dimer formation and thus have dominant-negative effects on the function of the remaining wild-type allele. The extreme variability in phenotypic expressivity and penetrance suggests that depending on the mutation's type and localisation, they have different effects on *TBX1* function, and that other, genetic co-factors play an important role.

Mutations in *LRP6*, a well-studied co-receptor in WNT/β-catenin-dependent signalling pathway, have a pleiotropic role in disease (figure 4 and online supplementary table S3).[34] Lrp6 mimics WNT activity in enhancing neural crest formation,[35] which gives rise to diverse cell lineages including cranioneural crest cells. *LRP6* mutations are associated with congenital neural tube defects,[36 37] and severe tooth agenesis,[38] sometimes with CLP.[39] The nonsense mutation we identified is located in the extracellular domain of *LRP6* causing haploinsufficiency. It is located adjacent to previously identified LoF heterozygous mutations, which cause autosomal tooth agenesis with reduced penetrance. In vitro studies showed that the p.Ala19Val mutant *LRP6* was retained in the endoplasmic reticulum,[38] which prevented it from being visible for incoming ligands at the cell surface. This suggests that the not fully penetrant tooth agenesis phenotype is due to loss of membrane-bound *LRP6* signalling. A combined phenotype (tooth agenesis and CLP) was observed in another patient with a truncating variant in the last exon of *LRP6* (intracellular; figure 4). Although it is expected to escape NMD, this variant is at least partially inactive since it loses three of its five signature PPPSP motifs,[40] crucial for phosphorylation of *LRP6* and *WNT* activation. Whether mutation type and position alone, or interaction with co-factors, explain the phenotypic variability associated with *LRP6* mutations remains to be seen.

GRHL3 was identified by linkage to the region 1p33–36 in a large Finnish family affected with VWS2 (MIM 606714). It was screened in a series of nsCP and VWS2 suspects. Mutations' effects ranged from dominant negative to strong and moderate hypomorphs (figure 5 and online supplementary table S4).[22 24 41] The latter mechanism loses some of its activity compared with the wild type, and as such predisposes rather than causes nsCP/VWS2. Our identified variant p.Arg391Cys has a dominant negative effect. It is associated with intriguing

variability in phenotypic expressivity. As a de novo mutation, it manifested with CLP and lip pits in one Filipino subject, while with isolated spina bifida in a second subject. We identified it as a germline mutation in two families of Belgian descent, both of which display a remarkable intrafamilial and interfamilial variability varying from PRS and CP to hypernasal speech (figure 1—F4a–F4b). Given that this mutation is a C>T transition, occurring in a CpG dinucleotide known to easily undergo deamination, this mutation likely represents a mutational 'hot-spot'. The occurrence in different ethnicities underscores this.

Lip pits, a hallmark of VWS, with an approximate penetrance of 80%, is observed in about 30% of patients with a *GRHL3* mutation.[22 41] However, lip pits were not present in either one of our families contrary to the aforementioned patient with VWS2. Although mild micrognathia was described, PRS observed in the proband and his sibling in family CLP-398 is novel and demonstrates once more the important clinical variability associated with variants in oral cleft 'syndrome-causing' genes, and the necessary quest to investigate modifiers.

Our results, as well as previously published data including *IRF6*,[20 42] *MSX1*,[43] *TBX22*[44] and *GRHL3*,[41] substantiate evidence for Mendelian transmission (with incomplete penetrance) of aetiological variants in 'syndrome-causing genes' in subclinical or non-syndromic cleft families. Therefore, custom-designed panels including CL/P-syndrome-causing genes should be used in diagnostic screens when a particular syndrome is suspected and also for families with nsCL/P. The benefits are dual; this will improve genetic counselling (by providing more informative chances of inheritance of 25%–50% versus an empiric of 4%–17%) and phenotyping of the patients (in search of minor/overlooked symptoms relevant to the syndrome) and also expand our understanding of molecular mechanisms underlying CL/P.

Such diagnostic screens will deliver many rare variants that will stratify CL/P cohorts more uniformly. Using genomic technologies such as WES or WGS (whole genome sequencing) on stratified cohorts will subsequently facilitate the teasing out of genetic modifiers such as coding variants in other genes, distantly acting enhancers, promoters and other conserved non-coding regions. Another unexplored explanation for sporadic CL/P are somatic mutations in developmentally relevant tissues in one of the known syndrome-linked genes. Proof of this concept has already elegantly been shown for various sporadically occurring vascular anomalies.[45 46] Unfortunately, targeted exomes and/or genomes of relevant tissues to CL/P are not that feasible to perform, until we determine the exact relevant tissues.

Concluding remarks

When syndromes manifest at the mild end of their spectra, they are often overlooked by clinicians. Our study clearly demonstrates the need to broaden criteria for diagnostic genetic screens to include patients with nsCL/P with a family history, as in 10% of them a causative Mendelian mutation could be identified. It could be that in an important percentage of the remaining cases, the same genes are involved due to an indel or copy number alteration (CNVs) not easily detectable by WES. WGS will therefore likely be the best choice for diagnostic screens of families with a history of nsCL/P.

Author affiliations
[1]Human Molecular Genetics, de Duve Institute, University of Louvain, Brussels, Belgium
[2]Center for Human Genetics, CLAD nord de France, CHU Amiens, Amiens, France
[3]EA 4666, Université Picardie Jules Verne, Amiens, France
[4]Center for Human Genetics, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium
[5]Pediatric Dentistry and Oral Care for Special Needs, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium
[6]Department of Orthodontics and Dentofacial Orthopedics, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium
[7]Service de génétique clinique Guy Fontaine, CHRU de Lille - Hôpital Jeanne de Flandre, Lille, France
[8]Service of Maxillofacial Surgery and Stomatology, CHU Amiens-Picardie, Amiens, France
[9]Department of Pediatrics, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium
[10]Division of Plastic Surgery, Centre Labio-Palatin, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

Contributors MB and MV prepared the manuscript. MB performed patient selection for WES, WES data analysis and validation/co-segregation of identified variants. BD, NR and BB were in charge of enrolment of subjects and collecting samples. RH helped in bioinformatic analyses of WES data. BD, NR, ST, SBS, OB, BD, GF and BB collected the clinical data. All authors contributed to re-drafting of the manuscript. MV conceived and coordinated the project and is responsible for the overall content. All authors have seen and approved the final manuscript.

Funding These studies were partially supported by funding from the Belgian Science Policy Office Interuniversity Attraction Poles (BELSPO-IAP) programme through the project IAP P7/43-BeMGI; BridgeIris RBC/2013-PFS-EH-11; the Fonds de la Recherche Scientifique—FNRS; CdR: J.0080.16 (all to MV). We also acknowledge the support of la Communauté française de Wallonie-Bruxelles, la Lotterie Nationale, Belgium and la région Haut de France, France. The authors thank the Genomics Platform of Université catholique de Louvain for next-generation sequencing analyses and the Foundation against Cancer, Belgium.

Competing interests None declared.

Patient consent Not required.

Ethics approval This study obtained 'La Commission d'Ethique Biomédicale Hospitalo-Facultaire' approval (2015/02NOV/572) by the medical faculty at University of Louvain, Brussels, Belgium.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All data relating to the study are available within the manuscript and accompanying online supplementary files.

REFERENCES

1. Mossey PA, Modell B. Epidemiology of oral clefts 2012: an international perspective. *Front Oral Biol* 2012;16:1–18.
2. Mulliken JB. The changing faces of children with cleft lip and palate. *N Engl J Med* 2004;351:745–7.
3. 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;32:285–9.
4. Curtis EJ. Congenital cleft lip and palate. *Am J Dis Child* 1961;102:853–7.
5. Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001;17:502–10.
6. Ghassibé M, Bayet B, Revencu N, Verellen-Dumoulin C, Gillerot Y, Vanwijck R, Vikkula M. Interferon regulatory factor-6: a gene predisposing to isolated cleft lip with or without cleft palate in the Belgian population. *Eur J Hum Genet* 2005;13:1239–42.
7. Ghassibe-Sabbagh M, Desmyter L, Langenberg T, Claes F, Boute O, Bayet B, Pellerin P, Hermans K, Backx L, Mansilla MA, Imoehl S, Nowak S, Ludwig KU, Baluardo C, Ferrian M, Mossey PA, Noethen M, Dewerchin M, François G, Revencu N, Vanwijck R, Hecht J, Mangold E, Murray J, Rubini M, Vermeesch JR, Poirel HA, Carmeliet P, Vikkula M. FAF1, a gene that is disrupted in cleft palate and has conserved function in zebrafish. *Am J Hum Genet* 2011;88:150–61.
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;41:473–7.
9. Grant SF, Wang K, Zhang H, Glaberson W, Annaiah K, Kim CE, Bradfield JP, Glessner JT, Thomas KA, Garris M, Frackelton EC, Otieno FG, Chiavacci RM, Nah HD, Kirschner RE, Hakonarson H. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. *J Pediatr* 2009;155:909–13.
10. Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T, Fallin MD, Redett RA, Raymond G, Schwender H, Jin SC, Cooper ME, Dunnwald M, Mansilla MA, Leslie E, Bullard S, Lidral AC, Moreno LM, Menezes R, Vieira AR, Petrin A, Wilcox AJ, Lie RT, Jabs EW, Wu-Chou YH, Chen PK, Wang H, Ye X, Huang S, Yeow V, Chong SS, Jee SH, Shi B, Christensen K, Melbye M, Doheny KF, Pugh EW, Ling H, Castilla EE, Czeizel AE, Ma L, Field LL, Brody L, Pangilinan F, Mills JL, Molloy AM, Kirke PN, Scott JM, Scott JM, Arcos-Burgos M, Scott AF. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nat Genet* 2010;42:525–9.
11. 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;42:24–6.
12. Sun Y, Huang Y, Yin A, Pan Y, Wang Y, Wang C, Du Y, Wang M, Lan F, Hu Z, Wang G, Jiang M, Ma J, Zhang X, Ma H, Ma J, Zhang W, Huang Q, Zhou Z, Ma L, Li Y, Jiang H, Xie L, Jiang Y, Shi B, Cheng J, Shen H, Wang L, Yang Y. Genome-wide association study identifies a new susceptibility locus for cleft lip with or without a cleft palate. *Nat Commun* 2015;6:6414.
13. Wolf ZT, Brand HA, Shaffer JR, Leslie EJ, Arzi B, Willet CE, Cox TC, McHenry T, Narayan N, Feingold E, Wang X, Sliskovic S, Karmi N, Safra N, Sanchez C, Deleyiannis FW, Murray JC, Wade CM, Marazita ML, Bannasch DL. Genome-wide association studies in dogs and humans identify ADAMTS20 as a risk variant for cleft lip and palate. *PLoS Genet* 2015;11:e1005059.
14. Leslie EJ, Carlson JC, Shaffer JR, Feingold E, Wehby G, Laurie CA, Jain D, Laurie CC, Doheny KF, McHenry T, Resick J, Sanchez C, Jacobs J, Emanuele B, Vieira AR, Neiswanger K, Lidral AC, Valencia-Ramirez LC, Lopez-Palacio AM, Valencia DR, Arcos-Burgos M, Czeizel AE, Field LL, Padilla CD, Cutiongco-de la Paz EM, Deleyiannis F, Christensen K, Munger RG, Lie RT, Wilcox A, Romitti PA, Castilla EE, Mereb JC, Poletta FA, Orioli IM, Carvalho FM, Hecht JT, Blanton SH, Buxó CJ, Butali A, Mossey PA, Adeyemo WL, James O, Braimah RO, Aregbesola BS, Eshete MA, Abate F, Koruyucu M, Seymen F, Ma L, de Salamanca JE, Weinberg SM, Moreno L, Murray JC, Marazita ML. A multi-ethnic genome-wide association study identifies novel loci for non-syndromic cleft lip with or without cleft palate on 2p24.2, 17q23 and 19q13. *Hum Mol Genet* 2016;25:ddw104–72.
15. Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, Becker J, Herberz R, AlChawa T, Nasser E, Böhmer AC, Mattheisen M, Alblas MA, Barth S, Kluck N, Lauster C, Braumann B, Reich RH, Hemprich A, Pötzsch S, Blaumeiser B, Daratsianos N, Kreusch T, Murray JC, Marazita ML, Ruczinski I, Scott AF, Beaty TH, Kramer FJ, Wienker TF, Steegers-Theunissen RP, Rubini M, Mossey PA, Hoffmann P, Lange C, Cichon S, Propping P, Knapp M, Nöthen MM. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. *Nat Genet* 2012;44:968–71.
16. Leslie EJ, Carlson JC, Shaffer JR, Butali A, Buxó CJ, Castilla EE, Christensen K, Deleyiannis FW, Leigh Field L, Hecht JT, Moreno L, Orioli IM, Padilla C, Vieira AR, Wehby GL, Feingold E, Weinberg SM, Murray JC, Beaty TH, Marazita ML. Genome-wide meta-analyses of nonsyndromic orofacial clefts identify novel associations between FOXE1 and all orofacial clefts, and TP63 and cleft lip with or without cleft palate. *Hum Genet* 2017;136:275–86.
17. Yu Y, Zuo X, He M, Gao J, Fu Y, Qin C, Meng L, Wang W, Song Y, Cheng Y, Zhou F, Chen G, Zheng X, Wang X, Liang B, Zhu Z, Fu X, Sheng Y, Hao J, Liu Z, Yan H, Mangold E, Ruczinski I, Liu J, Marazita ML, Ludwig KU, Beaty TH, Zhang X, Sun L, Bian Z. Genome-wide analyses of non-syndromic cleft lip with palate identify 14 novel loci and genetic heterogeneity. *Nat Commun* 2017;8:14364.
18. Beaty TH, Marazita ML, Leslie EJ. Genetic factors influencing risk to orofacial clefts: today's challenges and tomorrow's opportunities. *F1000Res* 2016;5:2800.
19. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–53.
20. Desmyter L, Ghassibe M, Revencu N, Boute O, Lees M, François G, Verellen-Dumoulin C, Sznajer Y, Moncla A, Benateau H, Claes K, Devriendt K, Mathieu M, Van Maldergem L, Addor MC, Drouin-Garraud V, Mortier G, Bouma M, Dieux-Coeslier A, Genevieve D, Goldenberg A, Gozu A, Makrythanasis P, McEntagart U, Sanchez A, Vilain C, Vermeer S, Connell F, Verheij J, Manouvrier S, Pierquin G, Odent S, Holder-Espinasse M, Vincent-Delorme C, Gillerot Y, Vanwijck R, Bayet B, Vikkula M. IRF6 screening of syndromic and a priori non-syndromic cleft lip and palate patients: identification of a new type of minor VWS sign. *Mol Syndromol* 2010;1:67–74.
21. Jugessur A, Shi M, Gjessing HK, Lie RT, Wilcox AJ, Weinberg CR, Christensen K, Boyles AL, Daack-Hirsch S, Trung TN, Bille C, Lidral AC, Murray JC. Genetic determinants of facial clefting: analysis of 357 candidate genes using two national cleft studies from Scandinavia. *PLoS One* 2009;4:e5385.
22. Peyrard-Janvid M, Leslie EJ, Kousa YA, Smith TL, Dunnwald M, Magnusson M, Lentz BA, Unneberg P, Fransson I, Koillinen HK, Rautio J, Pegelow M, Karsten A, Basel-Vanagaite L, Gordon W, Andersen B, Svensson T, Murray JC, Cornell RA, Kere J, Schutte BC. Dominant mutations in GRHL3 cause Van der Woude Syndrome and disrupt oral periderm development. *Am J Hum Genet* 2014;94:23–32.
23. Lemay P, De Marco P, Emond A, Spiegelman D, Dionne-Laporte A, Laurent S, Merello E, Accogli A, Rouleau GA, Capra V, Kibar Z. Rare deleterious variants in GRHL3 are associated with human spina bifida. *Hum Mutat* 2017;38:716–24.
24. Eshete MA, Liu H, Li M, Adeyemo WL, Gowans LJJ, Mossey PA, Busch T, Deressa W, Donkor P, Olaitan PB, Aregbesola BS, Braimah RO, Oseni GO, Oginni F, Audu R, Onwuamah C, James O, Augustine-Akpan E, Rahman LA, Ogunlewe MO, Arthur FKN, Bello SA, Agbenorku P, Twumasi P, Abate F, Hailu T, Demissie Y, Hailu A, Plange-Rhule G, Obiri-Yeboah S, Dunnwald MM, Gravem PE, Marazita ML, Adeyemo AA, Murray JC, Cornell RA, Butali A. Loss-of-function GRHL3 variants detected in African patients with isolated cleft palate. *J Dent Res* 2018;97:22034517729819.
25. Celli J, Duijf P, Hamel BC, Bamshad M, Kramer B, Smits AP, Newbury-Ecob R, Hennekam RC, Van Buggenhout G, van Haeringen A, Woods CG, van Essen AJ, de Waal R, Vriend G, Haber DA, Yang A, McKeon F, Brunner HG, van Bokhoven H. Heterozygous germline mutations in the p53 homolog p63 are the cause of EEC syndrome. *Cell* 1999;99:143–53.
26. Aradhya S, Lewis R, Bonaga T, Nwokekeh N, Stafford A, Boggs B, Hruska K, Smaoui N, Compton JG, Richard G, Suchy S. Exon-level array CGH in a large clinical cohort demonstrates increased sensitivity of diagnostic testing for Mendelian disorders. *Genet Med* 2012;14:594–603.
27. Vanbokhoven H, Melino G, Candi E, Declercq W. p63, a story of mice and men. *J Invest Dermatol* 2011;131:1196–207.
28. Jewell R, Eng B, Coates A, Hewitt S, Hobson E. 3q27.3 Microdeletion syndrome: further delineation of the second region of overlap and atopic dermatitis as a phenotypic feature. *Clin Dysmorphol* 2017;26:154–6.
29. Ponzi E, Asaro A, Orteschi D, Genuardi M, Zollino M, Gurrieri F. Variable expressivity of a familial 1.9 Mb microdeletion in 3q28 leading to haploinsufficiency of TP63: refinement of the critical region for a new microdeletion phenotype. *Eur J Med Genet* 2015;58:400–5.
30. Fulcoli FG, Franzese M, Liu X, Zhang Z, Angelini C, Baldini A. Rebalancing gene haploinsufficiency in vivo by targeting chromatin. *Nat Commun* 2016;7:11688.
31. Chen L, Fulcoli FG, Ferrentino R, Martucciello S, Illingworth EA, Baldini A. Transcriptional control in cardiac progenitors: Tbx1 interacts with the BAF chromatin remodeling complex and regulates Wnt5a. *PLoS Genet* 2012;8:e1002571.
32. Goodship J, Cross I, Scambler P, Burn J. Monozygotic twins with chromosome 22q11 deletion and discordant phenotype. *J Med Genet* 1995;32:746–8.
33. Pan Y, Wang ZG, Liu XY, Zhao H, Zhou N, Zheng GF, Qiu XB, Li RG, Yuan F, Shi HY, Hou XM, Yang YQ. A Novel TBX1 Loss-of-function mutation associated with congenital heart disease. *Pediatr Cardiol* 2015;36:1400–10.
34. Wang ZM, Luo JQ, Xu LY, Zhou HH, Zhang W. Harnessing low-density lipoprotein receptor protein 6 (LRP6) genetic variation and Wnt signaling for innovative diagnostics in complex diseases. *Pharmacogenomics J* 2017.
35. Tamai K, Semenov M, Kato Y, Spokony R, Liu C, Katsuyama Y, Hess F, Saint-Jeannet JP, He X. LDL-receptor-related proteins in Wnt signal transduction. *Nature* 2000;407:530–5.
36. Lei Y, Fathe K, McCartney D, Zhu H, Yang W, Ross ME, Shaw GM, Finnell RH. Rare LRP6 variants identified in spina bifida patients. *Hum Mutat* 2015;36:342–9.
37. Allache R, Lachance S, Guyot MC, De Marco P, Merello E, Justice MJ, Capra V, Kibar Z. Novel mutations in Lrp6 orthologs in mouse and human neural tube defects affect a highly dosage-sensitive Wnt non-canonical planar cell polarity pathway. *Hum Mol Genet* 2014;23:1687–99.
38. Massink MP, Créton MA, Spanevello F, Fennis WM, Cune MS, Savelberg SM, Nijman IJ, Maurice MM, van den Boogaard MJ, van Haaften G. Loss-of-function mutations in the WNT co-receptor LRP6 cause autosomal-dominant oligodontia. *Am J Hum Genet* 2015;97:621–6.
39. Ockeloen CW, Khandelwal KD, Dreesen K, Ludwig KU, Sullivan R, van Rooij I, Thonissen M, Swinnen S, Phan M, Conte F, Ishorst N, Gilissen C, RoaFuentes L, van de Vorst M, Henkes A, Steehouwer M, van Beusekom E, Bloemen M, Vankeirsbilck B, Bergé S, Hens G, Schoenaers J, Poorten VV, Roosenboom J, Verdonck A, Devriendt K, Roeleveldt N, Jhangiani SN, Vissers L, Lupski JR, de Ligt J, Von den Hoff JW, Pfundt R, Brunner HG, Zhou H, Dixon J, Mangold E, van Bokhoven H, Dixon MJ, Kleefstra T, Hoischen A, Carels CEL. Novel mutations in LRP6 highlight the role of WNT signaling in tooth agenesis. *Genet Med* 2016;18:1158–62.
40. Tamai K, Zeng X, Liu C, Zhang X, Harada Y, Chang Z, He X. A mechanism for Wnt coreceptor activation. *Mol Cell* 2004;13:149–56.
41. Mangold E, Böhmer AC, Ishorst N, Hoebel AK, Gültepe P, Schuenke H, Klamt J, Hofmann A, Gölz L, Raff R, Tessmann P, Nowak S, Reutter H, Hemprich A, Kreusch T, Kramer FJ, Braumann B, Reich R, Schmidt G, Jäger A, Reiter R, Brosch S, Stavusis J, Ishida M, Seselgyte R, Moore GE, Nöthen MM, Borck G, Aldhorae KA, Lace B, Stanier P, Knapp M, Ludwig KU. Sequencing the GRHL3 coding region reveals rare truncating

mutations and a common susceptibility variant for nonsyndromic cleft palate. *Am J Hum Genet* 2016;98:755–62.
42. Leslie EJ, Koboldt DC, Kang CJ, Ma L, Hecht JT, Wehby GL, Christensen K, Czeizel AE, Deleyiannis FW, Fulton RS, Wilson RK, Beaty TH, Schutte BC, Murray JC, Marazita ML. IRF6 mutation screening in non-syndromic orofacial clefting: analysis of 1521 families. *Clin Genet* 2016;90:28–34.
43. 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;40:399–407.
44. Marçano AC, Doudney K, Braybrook C, Squires R, Patton MA, Lees MM, Richieri-Costa A, Lidral AC, Murray JC, Moore GE, Stanier P. TBX22 mutations are a frequent cause of cleft palate. *J Med Genet* 2004;41:68–74.
45. Limaye N, Boon LM, Vikkula M. From germline towards somatic mutations in the pathophysiology of vascular anomalies. *Hum Mol Genet* 2009;18:R65–R74.
46. Limaye N, Kangas J, Mendola A, Godfraind C, Schlögel MJ, Helaers R, Eklund L, Boon LM, Vikkula M. Somatic activating PIK3CA mutations cause venous malformation. *Am J Hum Genet* 2015;97:914–21.

TP63	*List of variants detected in affected subjects*											
	(GRCh38/hg38)			***NM_003722.4***	***NP_003713.3***							
chr	***chr pos start***	***chr pos end***	***Change type***	***HGVS_Transcript consequence***	***HGVS_Protein consequence***	***Variant effect***	***Domains - Amino acids***	***Variant class***	***Phenotype***	***Reference***	***PMID***	***Allele freq gnomAD***	***Comments***
3	189631466	189822525	Gross DEL	ex. 1-4	?	del ex.1-4	TA 1-170	Patho.Var	EEC	Aradhya *et al* (2012)	22382802	Not Found	
3	189738658		SUB	c.208C>T	p.Gln70*	stop_gained	TA 1-170	Patho.Var	AEC	Eisenkraft *et al* (2015)	26470833	Not Found	Escapes NMD (translation re-initiation at next ATG)
3	189738739		SUB	c.289C>T	p.Arg97Cys	missense	TA 1-170	Patho.Var	SHFM	Zenteno *et al* (2005)	15736220	0.008137	
3	189808290		SUB	c.343G>T	p.Gly115Trp	missense	TA 1-170	Patho.Var	LMS	van Bokhoven *et al* (2002)	12037717	Not Found	
3	189808333		SUB	c.386C>T	p.Ser129Leu	missense	TA 1-170	VUS	nsCL	Leoyklang *et al* (2006)	16740912	0.004062	Mother is a carrier
3	189808333		SUB	c.386C>G	p.Ser129Trp	missense	TA 1-170	Patho.Var	LMS	Rinne *et al* (2007)	17224651	Not Found	
3	189808395		SUB	c.448G>A	p.Ala150Thr	missense	TA 1-170	Patho.Var	CP, hypohidrosis and oligodontia	Cabanillas *et al* (2011)	21906144	Not Found	
3	189808444		SUB	c.497C>T	p.Pro166Leu	missense	TA 1-170	Patho.Var	ADULT	van Zelst-Stams *et al* (2009)	19530185	Not Found	
3	189808465		SUB	c.518G>A	p.Gly173Asp	missense	DBD 170-362	Patho.Var	LMS	Rinne *et al* (2007)	17224651	Not Found	Additional phenotype: RHS Vera-Carbonell (2012) PMID:22342398
3	189808465		SUB	c.518G>T	p.Gly173Val	missense	DBD 170-362	Patho.Var	ADULT	Prontera *et al* (2011)	21990121	Not Found	
3	189864230		SUB	c.580-2A>C	p.Thr193_Tyr194insArg	in-frame (+1aa)	DBD 170-362	Patho.Var	SHFM	van Bokhoven *et al* (2001)	11462173	Not Found	Exon skipping experiment ->Causes insCCG for proline residue at amino acid 15, *(hg38 :p.Thr193_Tyr194insArg)*, leading to In-frame alternative-splicing product
3	189864230		SUB	c.580-2A>G	p.Thr193_Tyr194insArg	in-frame (+1aa)	DBD 170-362	Patho.Var	EEC	Monti *et al* (2013)	23463580	Not Found	Additional phenotype: SHFM+Amelogensis imperfecta Kantaputra et al (2012) PMID:22065540
3	189864250		SUB	c.598A>G	p.Lys200Glu	missense	DBD 170-362	Patho.Var	SHFM	Simonazzi *et al* (2012)	22430731	Not Found	
3	189864254		SUB	c.602T>C	p.Leu201Pro	missense	DBD 170-362	Patho.Var	EEC	Rinne *et al* (2006)	16691622	Not Found	
3	189864257		SUB	c.605A>G	p.Tyr202Cys	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2002)	12037717	Not Found	
3	189864343		SUB	c.691T>G	p.Tyr231Asp	missense	DBD 170-362	Patho.Var	EEC	Rinne *et al* (2007)	17224651	Not Found	
3	189864344		SUB	c.692A>G	p.Tyr231Cys	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2002)	12037717	Not Found	
3	189864346		SUB	c.694A>G	p.Lys232Glu	missense	DBD 170-362	Patho.Var	SHFM	van Bokhoven *et al* (2001)	11462173	Not Found	Additional phenotype: SHFM Wei et al (2012) PMID: 22574117
3	189864349		SUB	c.697A>G	p.Lys233Glu	missense	DBD 170-362	Patho.Var	SHFM	Ianakiev *et al* (2000)	10839977	Not Found	
3	189864373		SUB	c.721G>A	p.Val241Met	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2002)	12037717	Not Found	
3	189864379		SUB	c.727C>T	p.Arg243Trp	missense	DBD 170-362	Patho.Var	EEC	Celli *et al* (1999)	10535733	Not Found	Additional phenotype: ADULT/SHFM, Avitan-Hersh et al (2010) PMID: 21078104
3	189864380		SUB	c.728G>A	p.Arg243Gln	missense	DBD 170-362	Patho.Var	EEC	Celli *et al* (1999)	10535733	Not Found	Intermediate phenotype between ADULT and EEC , Otsuki et al (2016) PMID: 28293528
3	189864380		SUB	c.728G>T	p.Arg243Leu	missense	DBD 170-362	Patho.Var	EEC	de Mollerat *et al* (2003)	12525544	Not Found	
3	189864391		SUB	c.739C>G	p.His247Asp	missense	DBD 170-362	Patho.Var	EEC	Sorasio *et al* (2009)	19663851	Not Found	EEC without clefting
3	189864391		SUB	c.739C>T	p.His247Tyr	missense	DBD 170-362	Patho.Var	EEC	Rinne *et al* (2006)	16691622	Not Found	
3	189864392		SUB	c.740A>G	p.His247Arg	missense	DBD 170-362	Patho.Var	EEC	Clements *et al* (2010)	19903181	Not Found	
3	189866712		SUB	c.797G>A	p.Arg266Gln	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	Additional phenotype: SHFM, Su et al (2013) PMID: 23736768
3	189866712		SUB	c.797G>C	p.Arg266Pro	missense	DBD 170-362	Patho.Var	EEC	Sripathomsawat *et al* (2011)	21204238	Not Found	
3	189866732	189866735	INS	c.819_820dupCC	p.Gln274Profs*4	frame-shift	DBD 170-362	Patho.Var	nsCLP			Not Found	Identified in this study
3	189867873		SUB	c.923G>A	p.Cys308Tyr	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189867879		SUB	c.929G>C	p.Ser310Thr	missense	DBD 170-362	Patho.Var	EEC	Yin *et al* (2010)	20410354	Not Found	
3	189867882		SUB	c.932G>A	p.Ser311Asn	missense	DBD 170-362	Patho.Var	EEC	Celli *et al* (1999)	10535733	Not Found	
3	189867882		SUB	c.932G>C	p.Ser311Thr	missense	DBD 170-362	Patho.Var	EEC	Di Iorio *et al* (2012)	21959367	Not Found	
3	189867885		SUB	c.935G>A	p.Cys312Tyr	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2002)	12037717	Not Found	
3	189867902		SUB	c.952C>T	p.Arg318Cys	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189867902		SUB	c.952C>A	p.Arg318Ser	missense	DBD 170-362	Patho.Var	EEC	Clements *et al* (2010)	19903181	Not Found	
3	189867903		SUB	c.953G>A	p.Arg318His	missense	DBD 170-362	Patho.Var	EEC	Celli *et al* (1999)	10535733	Not Found	
3	189867903	189867904	SUB_INDEL	c.953_954delGCinsAA	p.Arg318Gln	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189867905		SUB	c.955C>T	p.Arg319Cys	missense	DBD 170-362	Patho.Var	SHFM/EEC	Ianakiev *et al* (2000)	10839977	Not Found	EEC (without CL/P) with linear and whorled naevoid hypermelanosis/EEC; Incomplete penetrance - mother carrier
3	189867905		SUB	c.955C>A	p.Arg319Ser	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189867906		SUB	c.956G>A	p.Arg319His	missense	DBD 170-362	Patho.Var	SHFM	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189867920	189867922	DEL	c.970_972delATT	p.Ile324del	in-frame (-1aa)	DBD 170-362	Patho.Var	EECUT plus syndrome	Giampietro *et al* (2013)	23613309	Not Found	
3	189868596		SUB	c.1009C>G	p.Arg337Gly	missense	DBD 170-362	Patho.Var	ADULT	Rinne *et al* (2006)	16691622	Not Found	
3	189868597		SUB	c.1010G>A	p.Arg337Gln	missense	DBD 170-362	Patho.Var	ADULT	Duijf *et al* (2002)	11929852	Not Found	
3	189868614		SUB	c.1027C>T	p.Arg343Trp	missense	DBD 170-362	Patho.Var	EEC	Celli *et al* (1999)	10535733	Not Found	Additional overlapping phenotype in a family RH, Brueggemann et al (2016) PMID: 26882220 , ADULT Chacon-camacho et al (2016) PMID :27028492

S1- List of TP63's rare variants detected in affected subjects

TP63	List of variants detected in affected subjects												
	(GRCh38/hg38)			NM_003722.4	NP_003713.3								
chr	chr pos start	chr pos end	Change type	HGVS_Transcript consequence	HGVS_Protein consequence	Variant effect	Domains - Amino acids	Variant class	Phenotype	Reference	PMID	Allele freq gnomAD	Comments
3	189868615		SUB	c.1028G>A	p.Arg343Gln	missense	DBD 170-362	Patho.Var	EEC	Ianakiev *et al* (2000)	10839977	Not Found	
3	189868615		SUB	c.1028G>C	p.Arg343Pro	missense	DBD 170-362	Patho.Var	EEC	de Molleratet *et al* (2003)	12525544	Not Found	
3	189868620		SUB	c.1033T>C	p.Cys345Arg	missense	DBD 170-362	Patho.Var	EEC	Celli *et al* (1999)	10535733	Not Found	
3	189868621		SUB	c.1034G>A	p.Cys345Tyr	missense	DBD 170-362	Patho.Var	EEC	Lehmann *et al* (2005)	15889277	Not Found	
3	189868624		SUB	c.1037C>A	p.Arg346Asp	missense	DBD 170-362	Patho.Var	EEC	Clements *et al* (2010)	19903181	Not Found	
3	189868624		SUB	c.1037C>G	p.Arg346Gly	missense	DBD 170-362	Patho.Var	EEC	Alves *et al* (2015)	25983622	Not Found	
3	189868626		SUB	c.1039T>A	p.Cys347Ser	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189868627		SUB	c.1040G>A	p.Cys347Tyr	missense	DBD 170-362	Patho.Var	AEC	Celik *et al* (2011)	22065614	Not Found	
3	189868629		SUB	c.1042C>T	p.Pro348Ser	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189868633		SUB	c.1046G>A	p.Gly349Glu	missense	DBD 170-362	Patho.Var	SHFM	Luo *et al* (2008)	17915261	Not Found	
3	189868635		SUB	c.1048A>G	p.Arg350Gly	missense	DBD 170-362	Patho.Var	EEC	Clements *et al* (2010)	19903181	Not Found	
3	189868638		SUB	c.1051G>A	p.Asp351Asn	missense	DBD 170-362	Patho.Var	EEC	Rinne *et al* (2006)	16691622	Not Found	
3	189868638		SUB	c.1051G>C	p.Asp351Asn	missense	DBD 170-362	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189868639		SUB	c.1052A>G	p.Asp351Gly	missense	DBD 170-362	Patho.Var	EEC	Akahoshi *et al* (2003)	12838557	Not Found	
3	189868640		SUB	c.1053C>A	p.Asp351Glu	missense	DBD 170-362	Patho.Var	EEC	Clements *et al* (2010)	19903181	Not Found	
3	189868641		SUB	c.1054A>G	p.Arg352Gly	missense	DBD 170-362	Patho.Var	nsCL	Leoyklang *et al* (2006)	16740912	Not Found	
3	189868648		SUB	c.1061C>A	p.Aal354Glu	missense	DBD 170-362	Patho.Var	EEC	Rinne *et al* (2007)	17224651	Not Found	
3	189886392		SUB	c.1350-2A>G	p.?	splice-site	SAM 541-607	Patho.Var	EEC	Barrow *et al* (2002)	12161593	Not Found	Predicted to induce a large splicing change intron<=30 - Table S4. Xiong et al (2015) PMID:25525159
3	189889478		SUB	c.1646T>C	p.Ile549Thr	missense	SAM 541-607	Patho.Var	RHS	Bertola *et al* (2004)	15200513	Not Found	
3	189890791		SUB	c.1655T>C	p.Phe552Ser	missense	SAM 541-607	Patho.Var	AEC	Berk *et al* (2009)	19840326	Not Found	
3	189890793		SUB	c.1657T>G	p.Leu553Val	missense	SAM 541-607	Patho.Var	AEC	McGrath *et al* (2001)	11159940	Not Found	
3	189890794		SUB	c.1658T>C	p.Leu553Ser	missense	SAM 541-607	Patho.Var	AEC	Payne *et al* (2005)	16365259	Not Found	
3	189890795		SUB	c.1659A>T	p.Leu553Phe	missense	SAM 541-607	Patho.Var	AEC	McGrath *et al* (2001)	11159940	Not Found	Mouse model. Defective expansion of ectodermal progenitor cells and impaired FGF signalling. Ferone et al (2012) PMID: 22247000
3	189890806		SUB	c.1670G>T	p.Gly557Val	missense	SAM 541-607	Patho.Var	AEC	Rinne *et al* (2007)	17224651	Not Found	
3	189890817		SUB	c.1681T>C	p.Cys561Arg	missense	SAM 541-607	Patho.Var	AEC	Chiu *et al* (2011)	19793345	Not Found	
3	189890817		SUB	c.1681T>G	p.Cys561Gly	missense	SAM 541-607	Patho.Var	AEC	McGrath *et al* (2001)	11159940	Not Found	
3	189890819		SUB	c.1683T>G	p.Cys561Trp	missense	SAM 541-607	Patho.Var	AEC	McGrath *et al* (2001)	11159940	Not Found	
3	189890821		SUB	c.1685T>C	p.Leu562Pro	missense	SAM 541-607	Patho.Var	AEC	Barbaro *et al* (2012)	22740388	Not Found	
3	189890825	189890826	INS	c.1689_1690insA	p.Tyr564Ilefs*14	frame-shift [mt&mz]	SAM 541-607	Patho.Var	EEC	Celli *et al* (1999)	10535733	Not Found	
3	189890829	189890830	DEL	c.1693_1694delTT	p.Phe565Hisfs*12	frame-shift [mt&mz]	SAM 541-607	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189890831		SUB	c.1695C>A	p.Phe565Leu	missense	SAM 541-607	Patho.Var	AEC	Rinne *et al* (2009)	19676060	Not Found	
3	189890842		SUB	c.1706G>T	p.Gly569Val	missense	SAM 541-607	Patho.Var	AEC	McGrath *et al* (2001)	11159940	Not Found	
3	189890850		SUB	c.1714A>C	p.Thr572Pro	missense	SAM 541-607	Patho.Var	AEC	McGrath *et al* (2001)	11159940	Not Found	
3	189890856	189890857	INS	c.1718_1720dupTCT	p.Ile573_Tyr574insPhe	in-frame (+1aa)	SAM 541-607	Patho.Var	AEC	Tsutsui *et al* (2003)	12932250	Not Found	
3	189890860		SUB	c.1724A>T	p.Gln575Leu	missense	SAM 541-607	Patho.Var	AEC	McGrath *et al* (2001)	11159940	Not Found	
3	189890863		SUB	c.1727T>C	p.Ile576Thr	missense	SAM 541-607	Patho.Var	AEC	McGrath *et al* (2001)	11159940	Not Found	
3	189890874		SUB	c.1738T>C	p.Ser580Pro	missense	SAM 541-607	Patho.Var	RHS	Kantaputra *et al* (2003)	12766194	Not Found	
3	189890875		SUB	c.1739C>T	p.Ser580Phe	missense	SAM 541-607	Patho.Var	AEC	Bertola *et al* (2004)	15200513	0.004064	
3	189890875		SUB	c.1739C>A	p.Ser580Tyr	missense	SAM 541-607	Patho.Var	RHS	Shotelersuk *et al* (2005)	15807690	Not Found	
3	189894206		SUB	c.1747G>T	p.Asp583Tyr	missense	SAM 541-607	Patho.Var	AEC	Beaudry *et al* (2009)	19353588	Not Found	[c.1747G>T]. Descr. as c.1630G>T. aka Asp544Tyr
3	189894206		SUB	c.1747G>T	p.Asp583Tyr	missense	SAM 541-607	Patho.Var	AEC	Rinne *et al* (2009)	19676060	Not Found	[c.1747G>T]. Descr. as c.1630G>T. aka Asp544Tyr
3	189894207		SUB	c.1748A>T	p.Asp583Val	missense	SAM 541-607	Patho.Var	AEC	Rinne *et al* (2009)	19676060	Not Found	[c.1748A>T]. Descr. as c.1631A>T. aka Asp544Val
3	189894210		SUB	c.1751T>C	p.Leu584Pro	missense	SAM 541-607	Patho.Var	AEC	Kannu *et al* (2006)	16532463	Not Found	
3	189894225		SUB	c.1766T>A	p.Ile589Asn	missense	SAM 541-607	Patho.Var	EEC	Al-Gazali *et al* (2010)	20437613	Not Found	
3	189894228		SUB	c.1769C>A	p.Pro590His	missense	SAM 541-607	Patho.Var	AEC	Hida *et al* (2014)	24898013	Not Found	
3	189894228		SUB	c.1769C>T	p.Pro590Leu	missense	SAM 541-607	Patho.Var	AEC	Rinne *et al* (2009)	19676060	Not Found	
3	189894240		SUB	c.1781G>C	p.Arg594Pro	missense	SAM 541-607	Patho.Var	AEC	Payne *et al* (2005)	16365259	Not Found	
3	189894249		SUB	c.1790T>C	p.Ile597Thr	missense	SAM 541-607	Patho.Var	AEC	Sawardekar *et al* (2011)	20738799	Not Found	
3	189894258		SUB	c.1799G>A	p.Gly600Asp	missense	SAM 541-607	Patho.Var	RHS	Steele *et al* (2005)	16190990	Not Found	
3	189894258		SUB	c.1799G>T	p.Gly600Val	missense	SAM 541-607	Patho.Var	AEC	Beaudry *et al* (2009)	19353588	Not Found	
3	189894264		SUB	c.1805T>C	p.Leu602Pro	missense	SAM 541-607	Patho.Var	EEC	Rinne *et al* (2006)	16691622	Not Found	
3	189894266		SUB	c.1807G>C	p.Asp603His	missense	SAM 541-607	VUS	nsCL	Leoyklang *et al* (2006)	16740912	0.04338	Also found in an unaffected father; found in 1/1000 control chromosomes.
3	189894274	189894274	DEL	c.1815delG	p.Gln606Serfs*98	frame-shift (stop lost) [mt&mz]	SAM 541-607	Patho.Var	RHS/AEC/ADULT	Serra *et al* (2011)	22069181	Not Found	Phenotype overlapping in one family
3	189894286	189894286	DEL	c.1827delA	p.Glu609Aspfs*95	frame-shift (stop lost) [mt&mz]	ITA 610-680	Patho.Var	RHS	Bougeard *et al* (2003)	12939657	Not Found	

TP63	*List of variants detected in affected subjects*												
	(GRCh38/hg38)			***NM_003722.4***	***NP_003713.3***								
chr	***chr pos start***	***chr pos end***	***Change type***	***HGVS_Transcript consequence***	***HGVS_Protein consequence***	***Variant effect***	***Domains - Amino acids***	***Variant class***	***Phenotype***	***Reference***	***PMID***	***Allele freq gnomAD***	***Comments***
3	189894297	189894297	DEL	c.1838delC	p.Pro613Leufs*91	frame-shift (stop lost) mt&mz	ITA 610-680	Patho.Var	RHS	Kannu *et al* (2006)	16532463	Not Found	
3	189894302	189894303	INS	c.1833_1843dupCTCCCCTTCTC	p.His615Profs*93	frame-shift (stop lost) mt&mz	ITA 610-680	Patho.Var	AEC / RHS	Prontera *et al* (2008)	19239083	Not Found	Intermediate phenotype between Hay-Wells and Rapp-Hodgkin syndromes in a patient
3	189894318	189894318	DEL	c.1859delC	p.Pro620Glnfs*84	frame-shift (stop lost) mt&mz	ITA 610-680	Patho.Var	AEC	van Bokhoven *et al* (2002)	12037717	Not Found	
3	189894319	189894320	DEL	c.1860_1861delAA	p.Ser621Glnfs*11	frame-shift (stop lost) mt&mz	ITA 610-680	Patho.Var	EEC	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189894359	189894359	DEL	c.1900delC	p.Arg634Glyfs*70	frame-shift (stop lost) mt&mz	ITA 610-680	Patho.Var	RHS	Holder-Espinasse *et al* (2007)	17609671	Not Found	
3	189894363		SUB	c.1904G>T	p.Gly635Val	missense	ITA 610-680	Patho.Var	ED	Goldsmith *et al* (2014)	24675753	Not Found	
3	189894363	189894363	DEL	c.1904delG	p.Gly635Valfs*69	frame-shift (stop lost) mt&mz	ITA 610-680	Patho.Var	RHS	Chan *et al* (2005)	15725251	Not Found	
3	189894369		SUB	c.1910G>T	p.Arg637Leu	missense	ITA 610-680	Patho.Var	AEC	Rinne *et al* (2009)	19676060	Not Found	
3	189894378		SUB	c.1919A>T	p.Asp640Val	missense	ITA 610-680	Patho.Var	AEC	Rinne *et al* (2009)	19676060	Not Found	
3	189894422	189894422	DEL	c.1963delC	p.Arg655Glufs*49	frame-shift (stop lost) mt&mz	ITA 610-680	Patho.Var	AEC	Rinne *et al* (2009)	19676060	Not Found	
3	189894433		SUB	c.1974G>A	p.Trp658*	stop_gained	ITA 610-680	Patho.Var	SHFM	Sowińska-Seidler *et al* (2014)	24163146	Not Found	
3	189894435	189894435	DEL	c.1976delA	p.Asn659Metfs*45	frame-shift (stop lost) mt&mz	ITA 610-680	Patho.Var	RHS	Dianzani *et al* (2003)	14684701	Not Found	
3	189894470		SUB	c.2011A>T	p.Lys671*	stop_gained	ITA 610-680	Patho.Var	LMS	Rinne *et al* (2006)	16691622	Not Found	
3	189894476		SUB	c.2017C>T	p.Gln673*	stop_gained	ITA 610-680	Patho.Var	SHFM	van Bokhoven *et al* (2001)	11462173	Not Found	
3	189894491		SUB	c.2032G>T	p.Glu678*	stop_gained	ITA 610-680	Patho.Var	SHFM	van Bokhoven *et al* (2002)	12037717	Not Found	

Legend:

Variant class

Patho.Var: Pathogenic Variant	All LoF and missense either neomutation, co-segregagting in >=2,or funct validated as deleterious
VUS: Variant of unknown significance	Rare >0.5% gnomAD
Assoc.Var: Associated Variant	>1% gnomAD frequent variant but losses some of it's activity compared to wild-type

Domains -Amino acids

TA	Transativation
DBD	DNA-binding domain
SAM	Sterile alpha motif
ITA	Inhibitor transactivation

Associated disorders

EEC	Ectrodactyly, ectodermal dysplasia, and cleft lip/palate syndrome 3
ADULT	acro-dermato-ungual-lacrimal-tooth
RHS	Rapp-Hodgkin syndrome
AEC	Ankyloblepharon-ectodermal defects-cleft lip/palate syndrome
LMS	Limb-mammary syndrome
SHFM	Split hand/foot malformation 1
nsCL	Non-syndromic cleft lip
nsCLP	Non-syndromic cleft lip and palate

Footnotes:

mt - stop codon predicted by mutationt@ster

mz - stop codon predicted by Mutalyzer

S1- List of TP63's rare variants detected in affected subjects

S1- List of TP63's rare variants detected in affected subjects

TBX1	*List of variants detected in affected subjects*												
	(GRCh38/hg38)			***NM_080647.1***	***NP_542378.1***								
chr	***chr pos start***	***chr pos end***	***Change type***	***HGVS_Transcript consequence***	***HGVS_Protein consequence***	***Variant effect***	***Domains - Amino acids***	***Variant class***	***Phenotype***	***Reference***	***PMID***	***Allele freq gnomAD***	***Comments***
22	19759558	19765825	Gross DEL	ex. 2-7	?	del ex.2-7		Patho.Var	TOF	Aguayo-Gómez *et al* (2015)	26036351	Not Found	De novo
22	19760999	19760055	Gross DEL	c.146_202del(57bp)	p.Arg49_Pro67del	deL 19 aa [mz]		Patho.Var	TOF	Griffin *et al* (2010)	20937753	Not Found	[c.146_202del] Descr.as c.129_185 (57bp), reduced transcriptional activity by 40%. Mother is a carrier
22	19761179	19761181	DEL	c.309_311delGAA	p.Lys103del	in-frame (-1aa) [mz]		Patho.Var	CHD	Xu *et al* (2017)	28272434	0.0004	Sig.reduced transcrptional activity and expression, parents DNA not available
22	19761255		SUB	c.385G>A	p.Glu129Lys	missense	T-Box 109-302	Patho.Var	CHD	Xu *et al* (2017)	24998776	Not Found	Sig. reduced transcriptional activity compared to Wt. (Fig. 2)
22	19763273		SUB	c.443T>A	p.Phe148Tyr	missense	T-Box 109-302	VUS	CTFA	Yagi *et al* (2003)	14585638	Not Found	
22	19764218		SUB	c.576C>T	p.Arg192Arg	synonymous_coding	T-Box 109-302	VUS	CHD	Gong *et al* (2001)	11748311	0.0449	
22	19764224		SUB	c.582C>G	p.His194Gln	missense	T-Box 109-302	Patho.Var	VCFS	Zweier *et al* (2007)	17273972	Not Found	
22	19765078		SUB	c.805C>T	p.Arg269*	stop-gained	T-Box 109-302	Patho.Var	VPI-isolated			Not Found	Identified in this study
22	19765102		SUB	c.829C>T	p.Gln277*	stop-gained	T-Box 109-302	Patho.Var	CHD	Pan *et al* (2015)	25860641	Not Found	
22	19765803		SUB	c.886C>A	p.Arg296Arg	synonymous_coding	T-Box 109-302	VUS	CHD	Gong *et al* (2001)	11748311	0.0481	
22	19765921		SUB	c.928G>A	p.Gly310Ser	missense			DGS	Yagi *et al* (2003)	14585638	0.0046	Unable to bind SMAD1 (Fig. 2), Fulcoli et al (2009) PubMed: 19557177; VCFS Soden et al (2014), PMID:25473035; Conotruncal heart defects, isolated ?; Homozygosis observed
22	19766428		SUB	c.1049G>A	p.Gly350Asp	missense		VUS	CHD	Gong *et al* (2001)	11748311	0.0004	
22	19766511		SUB	c.1132G>A	p.Gly378Ser	missense		VUS	MFDH	Simioni *et al* (2010)	20453311	0.0009	Father and daugher with a VSD
22	19766514	19766522	DEL	c.1135_1143delGCCGGCGGC	p.Ala379_Gly381del	in-frame (-3aa) [mz]		Patho.Var	CHD	Rauch *et al* (2004)	15060116	Not Found	
22	19766566		SUB	c.1187C>T	p.Pro396Leu	missense	NLS 394-448	VUS	CHD	Gong *et al* (2001)	11748311	0.0002	
22	19766602	19766602	DEL	c.1223delC	p.Ser408Trpfs*52	frame-shift [mz]		Patho.Var	CTFA	Yagi *et al* (2003)	14585638	Not Found	De novo
22	19766611		SUB	c.1232T>C	p.Leu411Pro	missense	NLS 394-448	Patho.Var	22q11.2 del features	Torres-Juan *et al* (2007)	17377518	Not Found	
22	19766632	19766632	DEL	c.1253delA	p.Tyr418Phefs*42	frame-shift [mz]		Patho.Var	22q11.2 del features	Ogata *et al* (2014)	24637876	Not Found	
22	19766653	19766660	DEL	c.1274_1281delACTATCTC	p.His425Argfs*189	frame-shift [mz]		Patho.Var	DGS	Gong *et al* (2001)	11748311	Not Found	[c.1321_1343del (23bp)] Descr. as c.1320_1342del (23bp)
22	19766700	19766721	Gross DEL	c.1321_1343del (23bp) [mz]	p.Leu441Alafs*168	frame-shift [mz]		Patho.Var	DGS/VCFS & psych disorder	Paylor *et al* (2006)	16684884	Not Found	
22	19766713	19766727	DEL	c.1334_1348del15	p.Gly445_His449del	in-frame (-5aa)		VUS	DGS	Gong *et al* (2001)	11748311	Not Found	
22	19766751	19766752	INS	c.1370_1372dupACC	p.His457dup	in-frame (+1aa)		VUS	VCFS	Gong *et al* (2001)	11748311	Not Found	[c.1399-1428dup (30bp)] Desc.r as c.1399-1427dup (30bp)
22	19766778	19766827	Gross Dup	c.1399-1428dup (30bp) [mz]	p.Ala467_Ala476dup	dup 10 aa		VUS	CHD	Gong *et al* (2001)	11748311	Not Found	
22	19766783	19766791	DEL	c.1404_1412delCGCTGCCGC	p.Ala474_Ala476del	in-frame (-3 aa) [mz]		VUS	DGS/VCFS	Gong *et al* (2001)	11748311	0.0359	

Legend:		
Variant class		
	Patho.Var: Pathogenic Variant	all LoF and missense either neomutation, co-segregagting in >=2,or funct validated as deleterious
	VUS: Variant of unknown significance	rare >0.5% gnomAD
	Assoc.Var: Associated Variant	>1% gnomAD frequent variant but losses some of it's activity compared to wild-type

Domains -Amino acids	
T-Box	DNA-binding domain
NLS	Nuclear Localisation Signal

Associated disorders	
MFDH	Midline facial defects with hypertelorism
CTFA	Conotruncal face anomaly syndrome
DGS	DiGeorge syndrome
VCFS	Velocardiofacial syndrome
CHD	Congenital heart defects
TOF	Tetralogy of Fallot
VPI	Velopharyngeal Insufficiency

Footnotes:

[mt] - stop codon predicted by mutationt@ster

[mz] - stop codon predicted by Mutalyzer

LRP6	*List of variants detected in affected subjects*													
	(GRCh38/hg38)			NM_002336.2	NP_002327.2									
chr	*chr pos start*	*chr pos end*	*Change type*	*HGVS_Transcript consequence*	*HGVS_Protein consequence*	*Variant effect*	*Domains (Amino acids)*	*Region*	*Variant class*	*Phenotype*	*Reference*	*PMID*	*Allele freq gnomAD*	*Comments*
12	12266728		SUB	c.8C>T	p.Ala3Val	missense	SP 1-24	*ECR*	Patho.Var	NTD	Lei *et al* (2015)	25546815	0.000009263	
12	12244655		SUB	c.56C>T	p.Ala19Val	missense	SP 1-24	*ECR*	Patho.Var	TA	Massink *et al* (2015)	26387593	0.000008574	
12	12244465		SUB	c.246A>T	p.Lys82Asn	missense		*ECR*	Patho.Var	CAD	Xu *et al* (2014)	24427284	0.000008237	
12	12203333		SUB	c.517C>G	p.Arg173Gly	missense	β-prop 125-258	*ECR*	VUS	TA	Ockeloen *et al* (2016)	26963285	Not Found	Segregation analysis not done; DNA N/A
12	12184040		SUB	c.916T>C	p.Tyr306His	missense	EGF 285-324	*ECR*	Patho.Var	NTD	Allache *et al* (2014)	24203697	0.00001648	
12	12181412		SUB	c.1004G>T	p.Arg335Leu	missense		*ECR*	VUS	nsCP	Ockeloen *et al* (2016)	26963285	Not Found	Inherited from unaffected mother; dent record N/A
12	12181337		SUB	c.1079G>A	p.Arg360His	missense	β-prop 352-565	*ECR*	Patho.Var	Mtb	Singh *et al* (2013)	23703864	0.0001318	
12	12181298		SUB	c.1118A>G	p.Tyr373Cys	missense	β-prop 352-565	*ECR*	Patho.Var	NTD	Allache et al (2014)	24203697	Not Found	
12	12181270	12181271	INS	c.1144_1145dupAG	p.Ala383Glyfs*8	frame-shift	β-prop 352-565	*ECR*	Patho.Var	TA	Massink *et al* (2015)	26387593	Not Found	
12	12181164		SUB	c.1252T>C	p.Tyr418His	missense	β-prop 352-565	*ECR*	Patho.Var	CAD	Guo *et al* (2016)	27455246	0.0003378	Coronary artery disease, normolipidaemic
12	12181118		SUB	c.1298A>G	p.Asn433Ser	missense	β-prop 352-565	*ECR*	Patho.Var	Mtb	Singh *et al* (2013)	23703864	0.00002471	
12	12179949		SUB	c.1406C>T	p.Pro469Leu	missense	β-prop 352-565	*ECR*	Patho.Var	TA	Ockeloen *et al* (2016)	26963285	Not Found	De novo
12	12179937		SUB	c.1418G>A	p.Arg473Gln	missense	β-prop 352-565	*ECR*	Patho.Var	Mtb	Singh *et al* (2013)	23703864	Not Found	
12	12179892		SUB	c.1463C>A	p.Ser488Tyr	missense	β-prop 352-565	*ECR*	Patho.Var	CAD	Xu *et al* (2014)	24427284	Not Found	
12	12165232		SUB	c.1609G>A	p.Gly537Arg	missense	β-prop 352-565	*ECR*	Patho.Var	TA	Ockeloen *et al* (2016)	26963285	Not Found	Full segregation in 3 generations
12	12165221		SUB	c.1620G>T	p.Leu540Phe	missense	β-prop 352-565	*ECR*	VUS	nsCL	Ockeloen *et al* (2016)	26963286	Not Found	Segregation analysis not done; DNA & dent records N/A
12	12165210		SUB	c.1631A>G	p.Tyr544Cys	missense	β-prop 352-565	*ECR*	Patho.Var	NTD	Lei *et al* (2015)	25546815	0.00001647	
12	12164545	12164546	INS	c.1779dupT	p.Glu594*	stop_gained	EGF 591-628	*ECR*	Patho.Var	TA	Massink *et al* (2015)	26387593	Not Found	
12	12164494		SUB	c.1831C>T	p.Arg611Cys	missense	EGF 591-628	*ECR*	Patho.Var	CAD	Mani *et al* (2007)	17332414	0.00001649	Proband homozygote p.Arg611Cys, other family memebers heterozygote for p.Arg611Cys and had an early onset of Coronary artery disease
12	12162414		SUB	c.2058C>G	p.Ile686Met	missense	β-prop 654-865	*ECR*	VUS	nsCL/P	Ockeloen *et al* (2016)	26963287	Not Found	Not maternally inhertited ; dent records N/A
12	12162246	12162247	INS	c.2224_2225dupTT	p.Leu742Phefs*7	frame-shift	β-prop 654-865	*ECR*	Patho.Var	TA	Massink *et al* (2015)	26387593	Not Found	
12	12158885	12158885	DEL	c.2735_2735delT	p.Phe912Leufs*37	frame-shift	EGF 892-930	*ECR*	Patho.Var	ASD	Iossifov *et al* (2014)	25363768	Not Found	
12	12150835		SUB	c.2994+1G>A	p.?	splice-site (donor)	β-prop 957-1177	*ECR*	Patho.Var	TA	Ockeloen *et al* (2016)	26963285	Not Found	Not maternally inhertited
12	12149153		SUB	c.2995G>C	p.Gly999Arg	missense	β-prop 957-1177	*ECR*	VUS	nsCL	Ockeloen *et al* (2016)	26963288	Not Found	Segregation analysis not done; DNA & dent records N/A
12	12148952		SUB	c.3196C>A	p.Pro1066Thr	missense	β-prop 957-1177	*ECR*	Patho.Var	CAD	Xu *et al* (2014)	24427284	Not Found	
12	12147398		SUB	c.3365A>C	p.Asp1122Ala	missense	β-prop 957-1177	*ECR*	VUS	nsCP	Ockeloen *et al* (2016)	26963289	Not Found	Not maternally inhertited; dent records N/A
12	12147390		SUB	c.3373C>T	p.Arg1125*	stop_gained	β-prop 957-1177	*ECR*	Patho.Var	TA+CLP			Not Found	Identified in this study ; dental records N/A
12	12138536		SUB	c.3398-2A>C	p.?	splice-site (acceptor)	β-prop 957-1177	*ECR*	Patho.Var	TA+CLP	Ockeloen *et al* (2016)	26963285	Not Found	Segregation analysis not done; DNA N/A
12	12135291		SUB	c.3617C>A	p.Pro1206His	missense	EGF 1206-1250	*ECR*	Patho.Var	CAD	Xu *et al* (2014)	24427284	Not Found	
12	12132001		SUB	c.3790A>G	p.Ile1264Val	missense	LDL	*ECR*	Patho.Var	CAD	Xu *et al* (2014)	24427284	0.000008247	
12	12126923		SUB	c.4082-2A>G	p.?	splice-site (acceptor)	LDL	*ECR*	Patho.Var	TA	Ockeloen *et al* (2016)	26963285	Not Found	Maternally inherited
12	12126867		SUB	c.4136G>A	p.Gly1379Asp	missense	LDL	*transmembrane*	VUS	TA+CLP	Ockeloen *et al* (2016)	26963289	Not Found	Inherited from unaffected father
12	12126847		SUB	c.4156G>T	p.Val1386Leu	missense	LDL	*transmembrane*	Patho.Var	NTD	Allache *et al* (2014)	24203697	Not Found	
12	12126705		SUB	c.4298C>T	p.Ser1433Leu	missense	1393-1613	*ICR*	VUS	TA+CLP	Ockeloen *et al* (2016)	26963289	Not Found	Inherited from unaffected father
12	12121375	12121375	DEL	c.4594delT	p.Cys1532Alafs*16	frame-shift	1393-1613	*ICR*	Patho.Var	TA+CLP	Ockeloen *et al* (2016)	26963285	Not Found	Inherited from affected mother
12	12121247		SUB	c.4721G>T	p.Arg1574Leu	missense	1393-1613	*ICR*	Patho.Var	NTD	Lei *et al* (2015)	25546815	Not Found	

Legend:

Variant class

Patho.Var: Pathogenic Variant	all LoF and missense either neomutation, co-segregagting in >=2,or funct validated as deleterious
VUS: Variant of unknown significance	rare >0.5% gnomAD
Assoc.Var: Associated Variant	>1% gnomAD frequent variant but losses some of it's activity compared to wild-type

Domains -Amino acids		**Region**	
SP	signal peptide	*ECR*	extracellular
β-prop	β propeller		
EGF	epidermal growth factor like domain		
LDL	LDL type A like receptor	Transmemebrane	1371-1393
		ICR	intracellular

Associated disorders

CAD	Coronary Artery Disease
Mtb	Metabolic syndrome
ASD	Autism Spectar Disorder
NTD	Neural Tube Defects
TA	Tooth Agenesis
TA+CLP	Combined tooth ageneis with cleft lip and palate
nsCL/P	Non -syndromic cleft lip (nsCL) with or without cleft palate (nsCP)

GRHL3	List of variants detected in affected subjects											
	(GRCh38/hg38)			NM_198174.2	NP_937817.3							
chr	chr pos start	chr pos end	Change type	HGVS_Transcript consequence	HGVS_Protein consequence	Variant effect	Domains - Amino acids	Variant class	Phenotype	Reference	PMID	Allele freq gnomAD	comments
1	24331456		INS	c.48dupC	p.Asp16Aspfs*10	frame-shift		Patho.Var	NTD	Lemay *et al* (2017)	28276201	Not Found	Consanguineous family: homozygous 1child (CLP+NTD+ scoliosis), homozygous 2nd child (NTD+ scoliosis), parents carriers
1	24336483	24336493	DEL	c.268_278del11	p.Tyr90Hisfs*4	frame-shift		Patho.Var	VWS2	Peyrard-Janvid *et al* (2014)	24360809	Not Found	Dominant negative -in an israeli family V-III, compound heterozygous
1	24336685	24336685	DEL	c.470delC	p.Gly158Alaf*r55	frame-shift		Patho.Var	nsCP	Eshete *et al* (2017)	28886269	Not Found	Strongly hypomorphic or null/parent carrier (African cohort)
1	24336712		SUB	c.497C>A	p.Pro166His	missense		VUS	nsCP	Eshete *et al* (2017)	28886269	Not Found	Modestly hypomorphic/parent carrier (African cohort)
1	24337687		SUB	c.738C>T	p.Gly246fs*10	frame-shift	DBD 230-423	Patho.Var	nsCP	Mangold *et al* (2016)	27018475	Not Found	De novo
1	24337790		SUB	c.840+1G>T	p.?	splice-site (donor)	DBD 230-423	Patho.Var	nsCP	Mangold *et al* (2016)	27018475	Not Found	De novo
1	24338044		SUB	c.893G>A	p.Arg298His	missense	DBD 230-423	Patho.Var	VWS2	Peyrard-Janvid *et al* (2014)	24360809	0.02447	Family also carries IRF6 c.239A>G.
1	24338067	24338068	INS	c.916dupC	p.Arg306Profs*11	frame-shift	DBD 230-423	Patho.Var	nsCP	Mangold *et al* (2016)	27018475	Not Found	Complete co-segregation
1	24339686	24339687	INS	c.971_972insGT	p.Phe324Leufs*22	frame-shift	DBD 230-423	Patho.Var	VWS2	Peyrard-Janvid *et al* (2014)	24360809	Not Found	Dominant negative
1	24339764		SUB	c.1047+2T>C	p.Ala318Glyfs*26	frame-shift	DBD 230-423	Patho.Var	NTD	Lemay *et al* (2017)	28276201	Not Found	Sporadic spina bifida: alternative splicing causes skipping exon 8
1	24342238		SUB	c.1171C>T	p.Arg391Cys	missense	DBD 230-423	Patho.Var	VWS2 or nsCP or NTD	Peyrard-Janvid *et al* (2014)	24360809	0.004096	Identified in this study as nsCP germline mutations in 2 families CLP-398 and CLP-986 and in other studies as de novo VWS2 and as de novo NTD
1	24342716		SUB	c.1229A>G	p.Asp410Gly	missense	DBD 230-423	VUS	nsCP	Eshete *et al* (2017)	28886269	0.004061	Strongly hypomorphic or null, parent carrier (African cohort)
1	24342769		SUB	c.1282A>C	p.Ser428Arg	missense		VUS	nsCP	Eshete *et al* (2017)	28886269	Not Found	Dominant negative, parent N/A (African cohort)
1	24342774	24342774	DEL	c.1285+2delT	p.?	splice-site (donor)		Patho.Var	nsCP	Mangold *et al* (2016)	27018475	Not Found	Complete co-segregation
1	24342967		SUB	c.1361C>T	p.Thr454Met	missense		Assoc. Var	nsCP	Leslie *et al* (2016)	27018472	0.02312	Disease-associated polymorphism with additional supporting functional evidence- retains 40% of its activity
1	24343026		SUB	c.1419+1G>T	p.?	splice-site (donor)		Patho.Var	VWS2	Peyrard-Janvid *et al* (2014)	24360809	Not Found	Dominant negative
1	24347483		SUB	c.1559G>A	p.Arg520Gln	missense	DD 493-602	Patho.Var	VWS2	Peyrard-Janvid *et al* (2014)	24360809	0.004061	Dominant negative
1	24347487	24347490	DEL	c.1563_1566delGGAG	p.Glu522Leufs*10	frame-shift	DD 493-602	Patho.Var	VWS2	Peyrard-Janvid *et al* (2014)	24360809	Not Found	Dominant negative
1	24347500	24347500	DEL	c.1576delG	p.Val526Cysfs*7	frame-shift	DD 493-602	Patho.Var	VWS2	Peyrard-Janvid *et al* (2014)	24360809	Not Found	Dominant negative
1	24347510		SUB	c.1586C>T	p.Ala529Val	missense	DD 493-602	Patho.Var	NTD	Lemay *et al* (2017)	28276201	0.06493	Sporadic spina bifida: reduced activation of target genes in a luciferase reprorter assay
1	24350089		SUB	c.1661A>G	p.Asn554Ser	missense	DD 493-602	Patho.Var	VWS2	Peyrard-Janvid *et al* (2014)	24360809	0.004065	Dominant negative -in an israeli family V-III, compound heterozygous
1	24350105		SUB	c.1677C>A	p.Tyr559*	stop_gained	DD 493-602	Patho.Var	nsCP	Eshete *et al* (2017)	28886269	Not Found	Modestly hypomorphic/parent carrier (African cohort)

Legend:

Variant class

Patho.Var: Pathogenic Variant	All LoF and missense either neomutation, co-segregagting in >=2,or funct validated as deleterious
VUS: Variant of unknown significance	Rare >0.5% gnomAD
Assoc.Var: Associated Variant	>1% gnomAD frequent variant but losses some of it's activity compared to wild-type

Domains -Amino acids

DBD	DNA-binding domain (230-423)
DD	dimerisation (493-602)

Associated disorders

NTD	Neural Tube Defects
VWS2	Van der Woude syndrome 2
nsCP	Non-syndromic cleft palate

B. Introductory remarks on publication II

"Unmasking familial CPX by WES and identification of novel clinical signs"

Bénédicte Demeer, Nicole Revencu, Raphael Helaers, Bernard Devauchelle, Geneviève Francois, Bénédicte Bayet , Miikka Vikkula

in American Journal of Medical Genetics

I selected the patients and performed the WES data analysis and validation/co-segregation of identified variants. Together with NR and BB, I clinically reassessed patients and families, and collected blood samples required for cosegregation studies.

This article shows how whole exome sequencing is an efficient tool to provide precise diagnosis in *a priori* non-syndromic CLP families, and how it could help to refine clinical entities.

UNMASKING FAMILIAL CPX BY WES AND IDENTIFICATION OF NOVEL CLINICAL SIGNS

Bénédicte Demeer[1,2,3], Nicole Revencu[1,4], Raphael Helaers[1], Bernard Devauchelle[3,5], Geneviève Francois[6], Bénédicte Bayet[7] , Miikka Vikkula[1]

[1] Human Molecular Genetics, de Duve Institute, University of Louvain, Brussels, Belgium

[2] Center for Human Genetics, CLAD nord de France, CHU Amiens, Amiens, France

[3] EA CHIMERE, Université Picardie Jules Verne, Amiens, France

[4] Center for Human Genetics, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

[5] Department of Maxillofacial Surgery and Stomatology, centre de compétence fentes et malformations faciales (MAFACE), CHU Amiens-Picardie, Amiens, France

[6] Department of Pediatrics, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

[7] Centre Labiopalatin, Division of Plastic Surgery, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

Correspondence to
Professor Miikka Vikkula,
Human Molecular Genetics
de Duve Institute, University of Louvain
Brussels 1348, Belgium
e-mail: miikka.vikkula@uclouvain.be
phone: + 32-2-764 7490 / fax: + 32-2-764 7460

KEYWORDS: *TBX22*, WES, non-syndromic CL/P, choanal atresia, Pierre-Robin sequence, reverse phenotyping.

ABSTRACT:

Mutations in the T-Box transcription factor gene *TBX22* are found in X-linked Cleft Palate with or without Ankyloglossia syndrome (CPX syndrome). In addition to X-linked inheritance, ankyloglossia, present in the majority of CPX patients, is an important diagnostic marker, but it is frequently missed or unreported, as it is a "minor" feature. Other described anomalies include cleft lip, micro and/or hypodontia, and features of CHARGE syndrome. We conducted Whole Exome sequencing (WES) on 22 individuals from 17 "a priori" non-syndromic cleft lip and/or cleft palate (CL/P) families. We filtered the data for heterozygous pathogenic variants within a set of predefined candidate genes. Two canonical splice-site mutations were found in *TBX22*. Detailed re-phenotyping of the 2 probands and their families unravelled orofacial features previously not associated with the CPX phenotypic spectrum: choanal atresia, Pierre-Robin sequence, and overgrowths on the posterior edge of the hard palate, on each side of the palatal midline. This study emphasizes the importance of WES analysis in familial CLP cases, combined with deep (reverse) phenotyping in "a priori" non-syndromic clefts.

INTRODUCTION

Orofacial clefts (OFCs) –including cleft lip with or without cleft palate (CL/P) and cleft palate (CP)- are among the most common birth defects, with an approximate incidence of 1/700 live births, depending on ethnicity, gender and cleft type (Mossey & Modell, 2012). Although not associated to a consequent rate of mortality in developed countries, they represent a significant lifelong morbidity, requiring multidisciplinary treatment and care. CL/P and CP have historically been considered as distinct disorders with separate etiologies based on the different embryological origin of the lip and the palate, and epidemiology (Fogh-Andersen, 1942; Fraser, 1955). However, both CL/P and CP can occur within the same family, suggesting some etiological overlap (Kondo et al., 2002). Approximately 30 % of clefts are syndromic, and in general follow a Mendelian-inheritance with an important variability in expressivity and incomplete penetrance. The remaining 70 % are non-syndromic (ns) (isolated), meaning that they do not occur in combination with other developmental, structural or cognitive anomalies. Isolated clefts were recognized to have an inherited component by studies of Fogh-Andersen (Fogh-Andersen, 1942). Non-syndromic clefts, of which 80% are sporadic and 20% are multiplex (familial) cases (Mossey & Modell, 2012), are considered to have a complex etiology with multiple predisposing genetic variants acting in concert with environmental factors.

Given the complex nature of non-syndromic clefts, identification of causal genes has been challenging. Several genetic approaches, often complementary, have been used to identify candidate regions and genes for nsCLP: identification of chromosomal abnormalities or microdeletions in individual cases, linkage analyses in families, direct sequencing, and candidate-locus and genome-wide association studies. Animal models, especially murine models, because of the similar development of their palate compared to humans, have provided candidate genes, and greatly helped in the understanding of gene function and dissection of the molecular pathways that underlie facial development.

Important findings have also emanated from studies involving syndromic forms. Some of those genes have now also been implicated in non-syndromic forms, through variable penetrance perhaps due to the action of modifiers (Schliekelman & Slatkin, 2002). Strongest evidence was obtained for the *IRF6* gene, causing van der Woude (VWS) syndrome, since the majority of the patients with VWS have only minor additional signs, and 15% have isolated CLP (Desmyter et al., 2010; Kondo et al., 2002). The second VWS gene, *GRHL3*, also demonstrated non-

complete penetrance for lip pits (Peyrard-Janvid et al., 2014). Mutations in *MSX1* are associated with syndromic clefting (with hypondontia) with approximately 2% of cases having non-syndromic CL/P; and *MSX1* polymorphisms are associated with non-syndromic CL in different populations. Several other syndromic genes, studied for their role in isolated CLP include *FGFR1, TP63, TBX22, PVRL1* and *PTCH*, and may contribute to up to 10% of non-syndromic CLP (Basha et al., 2018; Stanier & Moore, 2004)

Clefts of the lip and palate show a range of phenotypic expression. Yet, they are generally defined as "qualitative traits" (i.e.: affected versus unaffected). We conducted Whole Exome Sequencing (WES) on 22 individuals with "a priori" nsCLP drawn from multiplex families (n = 17) to identify causal nsCLP variants.

MATERIALS AND METHODS

PATIENT SELECTION

Samples for this study were drawn from a large cohort of nsCL/P & nsCP families, recruited over the last 25 years in collaboration with the multi-disciplinary team at the Centre Labio-Palatin (Cliniques universitaires Saint-Luc, UCL). Familial cases were defined as at least 2 affected cases within the third degree of relatedness. In order to collect samples uniformly, a standardized questionnaire was set up for the physician, notably detailing family history and clinical phenotype. A total of 22 affected patients were included, with index patients only for 12 families and an additional second or third degree relative for 5 families. Affected subjects were diagnosed as having: cleft lip alone: 9, cleft lip with cleft palate: 4, cleft palate alone: 5 or Pierre Robin sequence: 4. All patients had a standard karyotype and many had also been tested negative for *22q11.2* deletion and *IRF6* mutation. All participants gave informed consent for the participation in the study. The ethical committee of the medical faculty at University of Louvain, Brussels, Belgium, approved the research protocol.

WHOLE EXOME SEQUENCING (WES)

DNA was extracted from blood samples using Wizard® Genomic DNA Purification Kit (Promega). WES was carried out by using 1 µg of genomic DNA per sample. Exomes were captured using a commercial enrichment kit (Agilent SureSelectXT Human All Exon KitV5) and enriched libraries were sequenced on

an Illumina Hiseq2000. Generated reads were aligned to the human genome build hg19 using Burrows-Wheeler Aligner (BWA 0.7.15; see URLs). Data processing and variant calling were performed with Picard 1.107, Genome Analysis Toolkit 3.3 package (GATK) respectively. The generated variant (.vcf) files were imported into a database and further analysed using Highlander 14.10.3, an in-house bioinformatics framework for variant annotation, filtering, and visualization (see URLSs) (Helaers et al., under revision).

We considered as rare variants those with a frequency below 0.5% in public databases (ExAC, 1000 Genomes and GO-NL (genome of the Netherlands) and in our in-house control database.

As "likely pathogenic" variants, we considered those with a high impact (canonical splice-site variants and those generating a premature termination codon : out-of-frame indels and nonsense variants), or a moderate impact (non synonymous-NS), as annotated by SnpEff 4.1 and a Combined Annotation dependent Depletion (CADD) phred score > 10. Missense variants were considered as possibly pathogenic only if predicted to be possibly/probably damaging in at least three out of seven 7 variant in silico classifiers (Sift, Polyphen, LRT, Mutation taster, Mutation assessor, FATHMM, DEOGEN). We considered a list of 89 genes implicated in both syndromic and non-syndromic CLP (Stanier & Moore, 2004), nsCLP candidate genes (Leslie & Murray, 2013) and a selection of CLP candidate genes that we gathered from publications (Conte et al., 2016; Jugessur, Farlie, & Kilpatrick, 2009; Kousa, Mansour, Seada, Matoo, & Schutte, 2017; Lough, Byrd, Spitzer, & Williams, 2017) including recently confirmed CLP genes such as *ARHGAP29* (Leslie et al., 2012) and *GRHL3* (Peyrard-Janvid et al., 2014). Gene list is available upon request.

For all variants that satisfied our criteria, we performed co-segregation analysis on all the affected and non-affected family members that were available. Variants were PCR-amplified followed by sequencing on an ABI 3130XL genetic analyzer (Applied BioSystems).

RESULTS:

Using our filters and candidate gene list, we were able to point out two genetic variants in *TBX22* gene that we consider as causative of the respective phenotype (family 1 and 2). Within the other 15 families, no causative variant was detected. A nonsense variation in *PDGFC* gene (c.660G>A, NM_16205, p.Trp220*,

NP_057289) was found in the index case of one of the nsCP families, but did not cosegregate with the phenotype, and in one of the CL/P families, a missense variation was found in the *TULP4* gene (c.983A>G, NM_020245, p.Asn328Ser, NP_064630), but it did not cosegregate with the CL/P phenotype.

Family 1:

The proband was born at 40GW after an uneventful pregnancy, with birth weight of 2,910 kg. Diagnosis of a Pierre-Robin sequence (PRS) due to the association of cleft palate, glossoptosis and discrete micrognathia was made. Cerebral, cardiac and renal ultrasounds were normal. Array-CGH (Agilent ISCA 184) and *IRF6, COL2A1, COL11A1, COL11A2* sequencing were negative. On maternal side, familial history was remarkable with her mother having a cleft of soft palate, 1 male cousin a bifid uvula, 1 female cousin a cleft of soft and hard palate, 1 maternal great-uncle a cleft of soft and hard palate and the son of a maternal great-aunt a cleft of soft palate (Figure 1). Following genetic results, reassessment of the familial history noted 2 additional, more distant relatives with cleft palate: 1 male (II-1) and 1 female (III-13) (Figure 1). According to family members, II-11 had died at birth because of "closed nares". Clinical examination confirmed that the proband and her mother had no ankyloglossia, and that the maternal grandmother was unaffected.

We identified a G-to-A transversion at the +1 position of intron 2 (c.356+1G>A; NM_016954) of the *TBX22* gene in the proband. The mutation was heterozygous in the mother and the grandmother (Fig. 1). It was absent in the female cousin (Figure 1 – IV-8), a phenocopy with cleft of soft and hard palate. Further mutation segregation could not be performed in the extended family. This sequence change alters the invariant GT splice consensus sequence of exon 2 and is therefore predicted to disrupt normal splicing. It was not present in the single nucleotide polymorphism database (dbSNP), gnomAD or in control chromosomes (n =1128).

Family 2:

The proband is the second child of non-consanguineous parents, with a history of absent uvula on the maternal side. At birth, he presented with a short lingual frenulum that was cut during neonatal period. A diagnosis of submucous cleft

palate was made at 2,5 years of age. Following the genetic results, clinical reassessment of the family detected that his mother, who had a short frenulum cut soon after birth, and two more times later on, had two bony overgrowths on the posterior edge of the hard palate, on each side of the palatal midline, in addition to the absent uvula (Figure 2B). His maternal grandmother had a short lingual frenulum cut at the age of 15 years, and a surgical repair of a unilateral choanal atresia at 31 years of age. She also had two bony palatal overgrowths, and an underdeveloped soft palate. His sister had two bony overgrowths similar to her mother and maternal grandmother (Figure 2B). His maternal-great grandfather was described to have an absent uvula, associated with a short lingual frenulum, nasal speech and unilateral choanal atresia.

The proband and his maternal grandmother were found to have a G-to-A transversion at the +1 position of intron 4 (c.633+1G>A; NM_016954) in *TBX22*. Co-segregation studies confirmed that his mother was carrier, and found that his sister, who only had the bony overgrowths, was a carrier too. The substitution affects the invariant GT splice consensus sequence. It is the same variant reported already twice in classic CPX, once in a Brazilian (Braybrook et al., 2001) and once in a white European family (Marcano et al., 2004; Pauws et al., 2013).

DISCUSSION

We conducted a whole exome sequencing study on non-syndromic CLP families to search for rare causal variants. We identified two clear mutations in *TBX22* within the 17 non-syndromic CLP families screened. This result is consistent with the literature in favour of a 10% contribution of syndromic gene mutations to non-syndromic forms of clefting (Basha et al., 2018; Stanier & Moore, 2004).

Mutations in *TBX22* transcription factor are known to cause X-linked cleft palate (CPX) with ankyloglossia (MIM 303400) (Andreou et al., 2007; Braybrook et al., 2001; Braybrook et al., 2002; Chaabouni et al., 2005), as well as contributing to the general prevalence of orofacial clefting (Marcano et al., 2004; Suphapeetiporn, Tongkobpetch, Siriwan, & Shotelersuk, 2007). Patients with variable other anomalies including CL, micro and/or hypodontia (Kaewkhampa, Jotikasthira,

Malaivijitnond, & Kantaputra, 2012; Kantaputra et al., 2011), and features of CHARGE syndrome (Pauws et al., 2013) have also been reported. TBX22 linked CPX phenotype most likely results from a loss of protein function (Andreou et al., 2007; Braybrook et al., 2001). Other T-Box related diseases, such as ulnar mammary syndrome and Holt-Oram syndrome, have been linked to loss-of-function mutations in *TBX3* (Bamshad et al., 1997) and *TBX5* (Basson et al., 1997), respectively.

Canonical splice-site donor mutations affecting exon 4 (Braybrook et al., 2001; Marcano et al., 2004; Pauws et al., 2013) and exon 6 (Braybrook et al., 2001) have been associated with classical CPX phenotype. The c.356+1G>A mutation identified in family 1 is the first description of a canonical splice site donor mutation affecting exon 2. Like the previously described causal splice-site mutations, we predict that this variant would result either in aberrant splicing of exon 2, and/or in nonsense-mediated mRNA decay and haploinsufficiency. (Braybrook et al., 2001)

In classical CPX, a high inter- and intra-familial phenotypic variation exists, ranging from ankyloglossia alone to submucous cleft, bifid uvula or cleft of the secondary palate, all with or without ankyloglossia. With the X chromosomal localization, milder phenotype is usually exhibited by haploinsufficient females. Penetrance is complete in males and calculated to be 62% in females. Males frequently present cleft palate and ankyloglossia together (79%). While ankyloglossia is commonly seen (45%) in affected females, cleft palate (CP) is present as the sole feature in only 6% (Marcano et al., 2004).

The family 1 proband had isolated Pierre-Robin sequence (PRS), which has not been associated with classical CPX phenotype or to a *TBX22* mutation (a targeted PRS screen) (Marcano et al., 2004). Moreover, the limited pedigree size, the presence of a female phenocopy with CP, and family history of CP with similar expression in males and females prevent to define an X-linked inheritance pattern in this family. The absence of ankyloglossia, a useful diagnostic feature, underscores the difficulty to distinguish familial non syndromic CP from familial CPX. This patient and family do not *a priori* evoke need for screening *TBX22*.

The *TBX22* mutation is the most likely cause of the cleft/bifid uvula in family 1. As II-3, III-5, and IV-9 do not have ankyloglossia, the penetrance is lower than expected for CPX. Although minor signs could have been overlooked in the 5 unaffected obligate female carriers and 4 affected males, *TBX22* seems to have

retained some ”protective” activity or other (genetic) protective factors segregate in this family (Frischmeyer & Dietz, 1999). However IV-8 being a phenocopy and the unavailability of many affected members to be tested also raise the possibility of a more complicated cause or causes to run within the family.

In family 2, two clinical orofacial signs, previously not associated with CPX, were described: choanal atresia and bony overgrowths on the posterior edge of the hard palate, on each side of the palatal midline. Choanal atresia, a common craniofacial defect characterized by obstruction of the posterior nasal passages, has an approximate incidence of 1 case in 8.000 live births. Although the pathogenesis is still unknown, it is generally believed to be caused by the persistence of the buccopharyngeal membranes or by failure of the oronasal membrane to rupture. Additional malformations have been previously reported in up to 49% of individuals with choanal atresia and stenosis (CA/S) (Burrow et al., 2009; Keller & Kacker, 2000). Craniofacial expression of human and murine *TBX22* correlates with the cleft palate and ankyloglossia observed in patients (Braybrook et al., 2002). *Tbx22*^{null} mice have multiple craniofacial defects, which are all consistent with its known expression pattern, including submucous cleft palate with a small minority developing overt clefts, reduced vomer and ankyloglossia. *Tbx22*^{null} mice were also found to have persistent oro-nasal membrane or, in some mice, a partial rupture, resulting in choanal atresia (Pauws et al., 2009). This first description of choanal atresia in two CPX patients expands the orofacial phenotype associated with TBX22. Interestingly, in family 1, II-11 apparently died at birth because of “closed nares”, which evokes bilateral choanal stenosis. Specific investigation will be required in the future to determine the variability and penetrance of choanal atresia in TBX22.

In *Tbx22*^{null} mice, Pauws and co-workers showed that Tbx22 is a major determinant for intramembranous bone formation in the posterior hard palate, rather than being involved in palatal shelf closure. The knock-out mice had marked reduction in bone formation of the vomer and the posterior hard palate on analysis of the craniofacial skeleton. Interestingly, three of the affected patients of family 2 (Fig 2B) have two bony overgrowths on the posterior edge of the hard palate, on each side of the palatal midline, raising the question of an associated (reactive) abnormal ossification of the posterior palate. This has not been reported before, and it would be interesting to know if it is an overlooked or even specific *TBX22*-caused CPX sign. In family 2, I-1 and II-2 have a comparable, severe CPX phenotype. It

contrasts to the less severe phenotypic variation in generation III and IV implicating additional genetic and/or environmental factors to play a role.

Family recruitment in research studies is commonly based on the phenotype of the proband and a limited family history. This study provides support for "deep phenotyping" of probands and family members *a posteriori*. WES can guide the needed meticulous clinical evaluation regarding minor orofacial signs and endophenotypes, which may go unnoticed, as the least severe end of the spectrum of syndromic forms can mimic non-syndromic forms and a non-syndromic proband not always comes from a non-syndromic family. This study also highlights the contribution of "deep reverse phenotyping" to help refining clinical and molecular entities.

The recurrence risk is substantially enhanced for each a non-syndromic CLP patient with a causal mutation in a single-gene CLP-disorder, from the empiric 4-5% for non-syndromic CL/P and 3% for non-syndromic CP to the more personalized 25-50 %.

Substantial evidence is accumulating on rare variants in syndrome-causing genes to account for Mendelian transmission of both syndromic and non-syndromic familial clefting. Thus, one should consider to use NGS in clinics. This should not only be limited to syndromic CLP cases, but also used in familial cases of a priori non syndromic CLPs for genetic counselling indications. Two effective approaches could be used: targeted NGS (T-NGS) or WES. WES enables to investigate more genes throughout the exome, which might be relevant given the genetic heterogeneity of clefts. However, compared to T-NGS, it provides a weaker horizontal coverage, and it is more complex in terms of data analysis, less cost-effective, and raises the concern of incidental findings in genes not related to the phenotype (Christenhusz, Devriendt, & Dierickx, 2013). A candidate-gene strategy applied to WES would likely be the most relevant, with WES data being filtered through a fixed list of candidate genes, list that can evolve with biological knowledge. Clinical guidelines have still to be defined.

CONCLUDING REMARKS

Our study supports WES as an efficient tool for studying non-syndromic familial CLPs, as up to 10% of syndromic gene mutation may contribute to non-syndromic CLP. WES can thereby provide a precise diagnosis, accurate risk assessment and

help genetic counselling. It also emphasizes the important collaboration between clinicians and researchers to refine clinical entities and further improve our understanding of the genetic heterogeneity underlying non-syndromic CLP.

Contributors

BD and MV prepared the manuscript. BD performed patient selection for WES, WES data analysis and validation/co-segregation of identified variants. BD, NR and BB were in charge of enrolment of subjects and collecting samples. RH helped in bioinformatics analyses of WES data. BD, NR, BD, GF and BB collected the clinical data. All authors contributed to re-drafting of the manuscript. MV conceived and coordinated the project and is responsible for the overall content. All authors have seen and approved the final manuscript.

Funding

These studies were partially supported by funding from the Belgian Science Policy Office Interuniversity Attraction Poles (BELSPO-IAP) program through the project IAP P7/43-BeMGI; BridgeIris RBC/2013-PFS-EH-11; the Fonds de la Recherche Scientifique - FNRS; CdR: J.0080.16 (all to M.V.), and la Région Hauts-de-France, France (BD). We also acknowledge the support of la Communauté française de Wallonie-Bruxelles, and la Lotterie Nationale, Belgium.

Competing interests

None declared.

Acknowledgments

We are grateful to all the family members for their invaluable contributions. The authors thank the Genomics Platform of Université catholique de Louvain for Next Generation Sequencing, la Biobanque de Picardie -BB-0033-00017-, and the Foundation against Cancer, Belgium.

Web resources

Highlander: http://sites.uclouvain.be/highlander/

Online Mendelian Inheritance in Man (OMIM): http://www.omim.org

1000 Genomes: http://www.1000genomes.org/

Genome of the Netherlands (GoNL): http://www.nlgenome.nl/

NHLBI Exome Sequencing Project (ESP): http://evs.gs.washington.edu/EVS/

Exome Aggregation Consortium (ExAC): http://exac.broadinstitute.org/

Genome Aggregation Database (gnomAD): http://gnomad.broadinstitute.org/

The Human Gene Mutation Database (HMGD): http://www.hgmd.cf.ac.uk/

Polyphen2: http://genetics.bwh.harvard.edu/pph2/

SIFT: http://sift.jcvi.org/

Mutation Taster: http://www.mutationtaster.org/

Mutation Assessor: http://mutationassessor.org/

Functional Analysis through Hidden Markov Models (FATHMM): http://fathmm.biocompute.org.uk/

denovo-db: http://denovo-db.gs.washington.edu/denovo-db/

Mutalyzer: https://mutalyzer.nl/

Bibliography

Andreou, A. M., Pauws, E., Jones, M. C., Singh, M. K., Bussen, M., Doudney, K., . . . Stanier, P. (2007). TBX22 missense mutations found in patients with X-linked cleft palate affect DNA binding, sumoylation, and transcriptional repression. *Am J Hum Genet, 81*(4), 700-712. doi:10.1086/521033

Bamshad, M., Lin, R. C., Law, D. J., Watkins, W. C., Krakowiak, P. A., Moore, M. E., . . . Jorde, L. B. (1997). Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. *Nat Genet, 16*(3), 311-315. doi:10.1038/ng0797-311

Basha, M., Demeer, B., Revencu, N., Helaers, R., Theys, S., Bou Saba, S., . . . Vikkula, M. (2018). Whole exome sequencing identifies mutations in 10% of patients with familial non-syndromic cleft lip and/or palate in genes mutated in well-known syndromes. *J Med Genet*. doi:10.1136/jmedgenet-2017-105110

Basson, C. T., Bachinsky, D. R., Lin, R. C., Levi, T., Elkins, J. A., Soults, J., . . . Seidman, C. E. (1997). Mutations in human TBX5 [corrected] cause limb

and cardiac malformation in Holt-Oram syndrome. *Nat Genet, 15*(1), 30-35. doi:10.1038/ng0197-30

Braybrook, C., Doudney, K., Marcano, A. C., Arnason, A., Bjornsson, A., Patton, M. A., . . . Stanier, P. (2001). The T-box transcription factor gene TBX22 is mutated in X-linked cleft palate and ankyloglossia. *Nat Genet, 29*(2), 179-183. doi:10.1038/ng730

Braybrook, C., Lisgo, S., Doudney, K., Henderson, D., Marcano, A. C., Strachan, T., . . . Lindsay, S. (2002). Craniofacial expression of human and murine TBX22 correlates with the cleft palate and ankyloglossia phenotype observed in CPX patients. *Hum Mol Genet, 11*(22), 2793-2804.

Burrow, T. A., Saal, H. M., de Alarcon, A., Martin, L. J., Cotton, R. T., & Hopkin, R. J. (2009). Characterization of congenital anomalies in individuals with choanal atresia. *Arch Otolaryngol Head Neck Surg, 135*(6), 543-547. doi:10.1001/archoto.2009.53

Chaabouni, M., Smaoui, N., Benneji, N., M'Rad, R., Jemaa, L. B., Hachicha, S., & Chaabouni, H. (2005). Mutation analysis of TBX22 reveals new mutation in Tunisian CPX family. *Clin Dysmorphol, 14*(1), 23-25.

Christenhusz, G. M., Devriendt, K., & Dierickx, K. (2013). Disclosing incidental findings in genetics contexts: a review of the empirical ethical research. *Eur J Med Genet, 56*(10), 529-540. doi:10.1016/j.ejmg.2013.08.006

Conte, F., Oti, M., Dixon, J., Carels, C. E., Rubini, M., & Zhou, H. (2016). Systematic analysis of copy number variants of a large cohort of orofacial cleft patients identifies candidate genes for orofacial clefts. *Hum Genet, 135*(1), 41-59. doi:10.1007/s00439-015-1606-x

Desmyter, L., Ghassibe, M., Revencu, N., Boute, O., Lees, M., Francois, G., . . . Vikkula, M. (2010). IRF6 Screening of Syndromic and a priori Non-Syndromic Cleft Lip and Palate Patients: Identification of a New Type of Minor VWS Sign. *Mol Syndromol, 1*(2), 67-74. doi:10.1159/000313786

Fogh-Andersen, P. (1942). *Inheritance of harelip and cleft palate; contribution to the elucidation of the etiology of the congenital clefts of the face.* (Afhandling - Copenhagen), Copenhagen,.

Fraser, F. C. (1955). Thoughts on the etiology of clefts of the palate and lip. *Acta Genet Stat Med, 5*(4), 358-369.

Frischmeyer, P. A., & Dietz, H. C. (1999). Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet, 8*(10), 1893-1900.

Jugessur, A., Farlie, P. G., & Kilpatrick, N. (2009). The genetics of isolated orofacial clefts: from genotypes to subphenotypes. *Oral Dis, 15*(7), 437-453. doi:10.1111/j.1601-0825.2009.01577.x

Kaewkhampa, A., Jotikasthira, D., Malaivijitnond, S., & Kantaputra, P. (2012). TBX22 mutation associated with cleft lip/palate, hypodontia, and limb anomaly. *Cleft Palate Craniofac J, 49*(2), 240-244. doi:10.1597/10-208

Kantaputra, P. N., Paramee, M., Kaewkhampa, A., Hoshino, A., Lees, M., McEntagart, M., . . . Stanier, P. (2011). Cleft lip with cleft palate,

ankyloglossia, and hypodontia are associated with TBX22 mutations. *J Dent Res, 90*(4), 450-455. doi:10.1177/0022034510391052

Keller, J. L., & Kacker, A. (2000). Choanal atresia, CHARGE association, and congenital nasal stenosis. *Otolaryngol Clin North Am, 33*(6), 1343-1351, viii.

Kondo, S., Schutte, B. C., Richardson, R. J., Bjork, B. C., Knight, A. S., Watanabe, Y., . . . Murray, J. C. (2002). Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. *Nat Genet, 32*(2), 285-289. doi:10.1038/ng985

Kousa, Y. A., Mansour, T. A., Seada, H., Matoo, S., & Schutte, B. C. (2017). Shared molecular networks in orofacial and neural tube development. *Birth Defects Res, 109*(2), 169-179. doi:10.1002/bdra.23598

Leslie, E. J., Mansilla, M. A., Biggs, L. C., Schuette, K., Bullard, S., Cooper, M., . . . Murray, J. C. (2012). Expression and mutation analyses implicate ARHGAP29 as the etiologic gene for the cleft lip with or without cleft palate locus identified by genome-wide association on chromosome 1p22. *Birth Defects Res A Clin Mol Teratol, 94*(11), 934-942. doi:10.1002/bdra.23076

Leslie, E. J., & Murray, J. C. (2013). Evaluating rare coding variants as contributing causes to non-syndromic cleft lip and palate. *Clin Genet, 84*(5), 496-500. doi:10.1111/cge.12018

Lough, K. J., Byrd, K. M., Spitzer, D. C., & Williams, S. E. (2017). Closing the Gap: Mouse Models to Study Adhesion in Secondary Palatogenesis. *J Dent Res, 96*(11), 1210-1220. doi:10.1177/0022034517726284

Marcano, A. C., Doudney, K., Braybrook, C., Squires, R., Patton, M. A., Lees, M. M., . . . Stanier, P. (2004). TBX22 mutations are a frequent cause of cleft palate. *J Med Genet, 41*(1), 68-74.

Mossey, P. A., & Modell, B. (2012). Epidemiology of oral clefts 2012: an international perspective. *Front Oral Biol, 16*, 1-18. doi:10.1159/000337464

Pauws, E., Hoshino, A., Bentley, L., Prajapati, S., Keller, C., Hammond, P., . . . Stanier, P. (2009). Tbx22null mice have a submucous cleft palate due to reduced palatal bone formation and also display ankyloglossia and choanal atresia phenotypes. *Hum Mol Genet, 18*(21), 4171-4179. doi:10.1093/hmg/ddp368

Pauws, E., Peskett, E., Boissin, C., Hoshino, A., Mengrelis, K., Carta, E., . . . Stanier, P. (2013). X-linked CHARGE-like Abruzzo-Erickson syndrome and classic cleft palate with ankyloglossia result from TBX22 splicing mutations. *Clin Genet, 83*(4), 352-358. doi:10.1111/j.1399-0004.2012.01930.x

Peyrard-Janvid, M., Leslie, E. J., Kousa, Y. A., Smith, T. L., Dunnwald, M., Magnusson, M., . . . Schutte, B. C. (2014). Dominant mutations in GRHL3

cause Van der Woude Syndrome and disrupt oral periderm development. *Am J Hum Genet, 94*(1), 23-32. doi:10.1016/j.ajhg.2013.11.009

Schliekelman, P., & Slatkin, M. (2002). Multiplex relative risk and estimation of the number of loci underlying an inherited disease. *Am J Hum Genet, 71*(6), 1369-1385. doi:10.1086/344779

Stanier, P., & Moore, G. E. (2004). Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. *Hum Mol Genet, 13 Spec No 1*, R73-81. doi:10.1093/hmg/ddh052

Suphapeetiporn, K., Tongkobpetch, S., Siriwan, P., & Shotelersuk, V. (2007). TBX22 mutations are a frequent cause of non-syndromic cleft palate in the Thai population. *Clin Genet, 72*(5), 478-483. doi:10.1111/j.1399-0004.2007.00891.x

Figure 1 – Pedigree of family 1

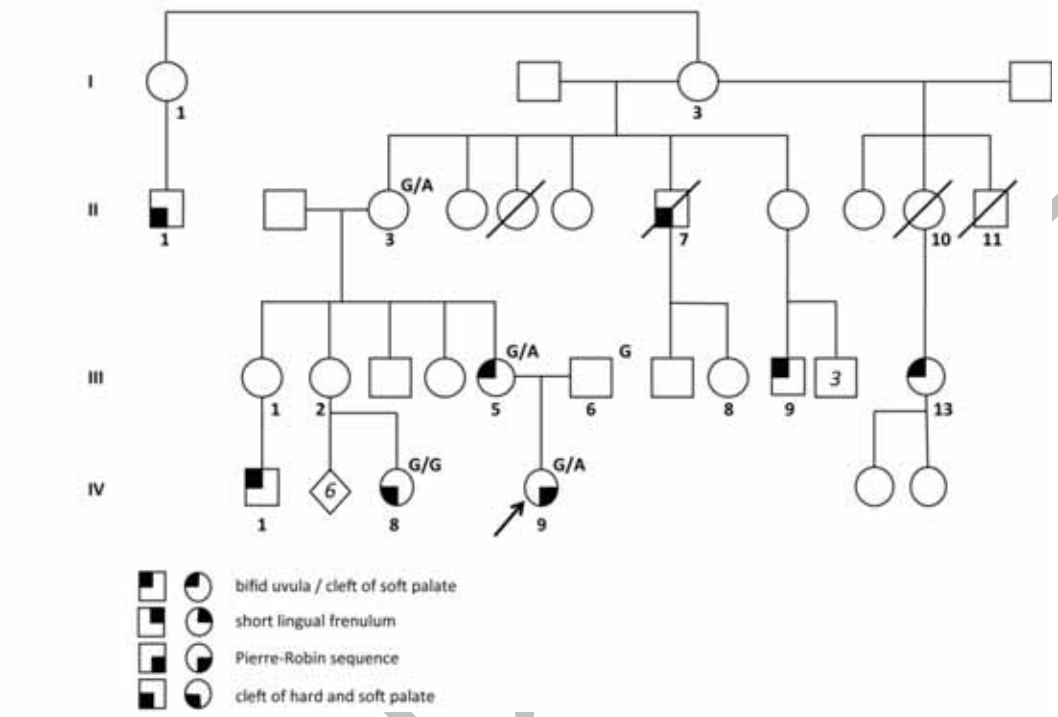
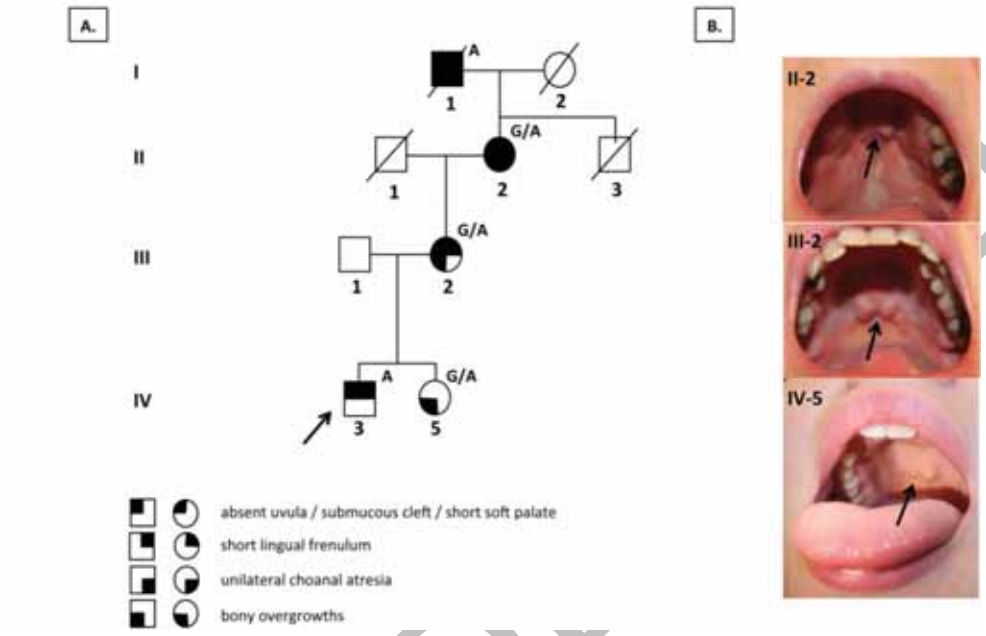


Figure 2 – A) Family 2 pedigree; B) Endobuccal pictures of II-2, III-2, IV-5. Arrows show bony overgrowths



C. Introductory remarks on publication III

"Deleterious mutations in one third of non-syndromic discontinuous cleft lip and palate patients"

Bénédicte Demeer, Nicole Revencu, Raphael Helaers, Cica Gbaguidi, Geneviève Francois, Bernard Devauchelle, Bénédicte Bayet, Miikka Vikkula

Submitted to Genetics in medicine

I focused on a specific CLP phenotype named "discontinuous cleft". I could find some cases in the laboratory cohort of 1200 pedigrees and could collect more cases thanks to BB and CG. I selected the patients and performed the WES data analysis and validation/co-segregation of identified variants. Together with NR, BB and CG, I clinically reassessed patients and families, and collected blood samples required for cosegregation studies.

This study performed on clinically well-defined subgroup of CLP appears to be a promising approach, as we could find causal mutations in a significant number of patients (3/8). We currently collect more Introductory remarks on preliminary result.

DELETERIOUS MUTATIONS IN ONE THIRD OF NON-SYNDROMIC DISCONTINUOUS CLEFT LIP AND PALATE PATIENTS

Bénédicte Demeer[1,2,3], Nicole Revencu[1,4], Raphael Helaers[1], Cica Gbaguidi [5], Geneviève Francois[6], Bernard Devauchelle[3,5], Bénédicte Bayet[7] , Miikka Vikkula[1]

[1] Human Molecular Genetics, de Duve Institute, University of Louvain, Brussels, Belgium

[2] Center for Human Genetics, CLAD nord de France, CHU Amiens, Amiens, France

[3] EA CHIMERE, Université Picardie Jules Verne, Amiens, France

[4] Center for Human Genetics, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

[5] Department of Maxillofacial Surgery and Stomatology, centre de compétence fentes et malformations faciales (MAFACE), CHU Amiens-Picardie, Amiens, France

[6] Department of Pediatrics, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

[7] Centre Labiopalatin, Division of Plastic Surgery, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

Correspondence to
Professor Miikka Vikkula,
Human Molecular Genetics
de Duve Institute, University of Louvain
Brussels, Belgium
e-mail: miikka.vikkula@uclouvain.be
phone: + 32-2-764 7490 / fax: + 32-2-764 7460

KEYWORDS: discontinuous CLP, non-syndromic CLP, *ARHGAP29, FGFR1, DLG1*, WES

Word count: 2981

ABSTRACT:

Oral clefts are composed of cleft of the lip, cleft of the lip and palate, or cleft of the palate, and they are associated with a wide range of expression and severity. When cleft of the palate is associated with a cleft of the lip with preservation of the primary palate, it defines an atypical phenotype called discontinuous cleft. Although this phenotype may represent 5% of clefts of the lip and/or palate (CLP), it is rarely specifically referred to and its pathophysiology is unknown. We conducted Whole Exome sequencing (WES) on 8 non-syndromic discontinuous CLP individuals in order to identify genes and mutations that could underlie this phenotype. We discovered loss-of-function mutations in 3 out of the 8 individuals implicating *FGFR1*, *ARGHAP29* and *DLG1* genes, representing almost 40% of this small cohort.

INTRODUCTION

Cleft of the lip and/or palate (CLP) are among the most common birth defects, with an approximate incidence of 1/700 live births and with a wide variability of expression depending on ethnicity, gender and cleft type. Although CLPs are not associated with a consequent rate of mortality in developed countries, they represent a significant lifelong morbidity. This requires multidisciplinary treatment and care.

CLPs are classified as syndromic and non-syndromic (isolated). The latter is not associated with additional developmental, structural and/or cognitive anomalies. Syndromic CLPs, which represent 30% of all CLPs, generally follow a Mendelian inheritance with an important variability in expressivity and incomplete penetrance. The other 70% of CLPs occur as isolated lesions, and are considered to have a complex etiology with both genetic and environmental factors acting in concert.

Given the complex and heterogeneous nature of non-syndromic cleft, different genetic approaches have been employed successfully to identify chromosomal loci and genes [13].

With rapid improvement in technology, in particular invent of next-generation sequencing and bioinformatic algorithms, we now have the ability to interrogate the entire exome or genome in order to identify risk loci [83]. Whole exome sequencing (WES) is successfully performed for Mendelian disorders, becoming a useful approach for proving clinical molecular diagnosis [200].This strategy has also been applied to identify the causes of complex traits including cleft lip and palate [183, 201].

Historically, CLP has been divided into cleft of the lip (CL), CL with cleft of the palate (CP), which are often co-classified as cleft of the lip with or without cleft palate (CL/P), and CP. This historical broad subdivision is based on a number of epidemiological features of each condition, and a distinct developmental origin of the primary palate and the secondary palate. Segregation of CL/P and CP has been exceptionally reported for families with etiologic mutations in specific genes: P63, MSX1, IRF6, and FGFR1 [13]. Clefts of the lip and palate show a range of phenotypic expression, and dividing CLP in a simplistic way to CL/P and CP has the potential to lose important information. Analysing separately CLs from the rest of the CL/P group in non-syndromic multiplex families revealed differences in linkage results [35]. It highlights the importance of subphenotyping, and requirement of carefully stratified cohorts to further understand the genetic heterogeneity underlying non-syndromic CLP (nsCLP).

One overlooked subphenotype within CL/Ps is the "discontinuous" clefting, which associates an incomplete cleft of the primary palate with cleft of the secondary palate. Concerning the description of this CLP, some authors use the term "primary palate" to refer to the anterior intermaxillary palate, whereas others used it to describe the tissues formed by the fusion between the maxillary and medial nasal processes : lip, alveolus and anterior intermaxillary palate [4]. We use the term discontinuous cleft for any cleft of the primary palate respecting part of the intermaxillary palate, associated with a cleft of the secondary palate. This atypical phenotype has been reported in the literature with diverse terminology, referred to as "separate cleft of the lip and palate", "unconnected cleft", "bridged cleft" and "interrupted cleft" [202, 203]. It has been seldomly separated from other subtypes, and likely, if not specified, included either in the CL or CL/P group. A recent report on 356 Brazilian individuals with discontinuous cleft estimated that this group represents 5% of CLP [202]. The pathophysiology behind this subgroup is currently unknown. In this study, we conducted WES on 8 individuals with non-syndromic discontinuous CLP in order to identify genes and mutations that contribute to this phenotype.

MATERIALS AND METHODS

PATIENT SELECTION

Clinical data and samples from patients with nsCLP and their family members were collected in collaboration with the Centre labiopalatin, Cliniques universitaires St Luc, Brussels, Belgium and Amiens-Picardie Hospital, France. Informed consent was signed by each participant prior to conducting the study, as approved by the institutional review board. Each patient was asked to fill in a questionnaire recapitulating personal, familial history and his/her malformations. The CLP phenotyping was based on clinical examination by the plastic surgeon while evaluating the infant before surgery and follow up, and on the questionnaire. Eight patients who met the clinical diagnosis of ns discontinuous cleft were selected and included in the study. This comprises 6 sporadic cases and 2 with a familial history of CL/P (Table 1).

WHOLE EXOME SEQUENCING (WES)

DNA was extracted from blood samples using Wizard® Genomic DNA Purification Kit (Promega). WES was carried out by using 1 µg of genomic DNA

per sample. Exomes were captured using a commercial enrichment kit (Agilent SureSelectXT Human All Exon KitV5) and enriched libraries were sequenced on an Illumina Hiseq2000. Generated reads were aligned to the human genome build hg19 using Burrows-Wheeler Aligner (BWA 0.7.15; see URLs). Data processing and variant calling were performed with Picard 1.107, Genome Analysis Toolkit 3.3 package (GATK) respectively. The generated variant (.vcf) files were imported into a database and further analysed using Highlander 14.10.3, an in-house bioinformatic framework for variant annotation, filtering, and visualization (see URLSs) (Helaers et al., under revision).

We considered as rare variants those with a frequency below 0.5% in public databases (ExAC, 1000 Genomes and GO-NL (genome of the Netherlands) and in our in-house control database. As "likely pathogenic" variants, we considered those with a high impact (canonical splice-site variants and those generating a premature termination codon=PTC (out-of-frame indels and nonsense variants)), or a moderate impact (non synonymous-NS), as annotated by SnpEff 4.1 and a Combined Annotation dependent Depletion (CADD) phred score > 15. Missense variants were considered as possibly pathogenic only if predicted to be possibly/probably damaging in at least three out of seven in silico variant classifiers (Sift, Polyphen, LRT, Mutation taster, Mutation assessor, FATHMM, DEOGEN). We considered a list of 361 genes previously implicated in any form of clefting as candidates [184].

For all variants that satisfied our criteria, we performed co-segregation analysis on all the affected and non-affected family members that were available. Variants were PCR-amplified followed by sequencing on an ABI 3130XL genetic analyzer (Applied BioSystems).

RESULTS:

Using our filters and candidate gene list, we were able to point out three genetic variants that we consider as causative of the respective phenotype in a mendelian way. These three mutations generated stop codons in genes that have already been linked to a human CLP phenotype and/or syndrome.

Patient 1

The proband is the only child of an unrelated Caucasian couple, with unremarkable family history. She was born at 37 WG with BW: 3,070kg and BH: 47 cm. She presented a right labioalveolar cleft lip associated to a submucous cleft. At 11 years

and 3 months of age, she has had no specific follow-up, and thus undetermined reproductive axis and olfactory clinical status.
We identified a 2 bp insertion (c.1809_1810insTC: NM_001174067) in the *FGFR1* gene. This mutation was not found in her mother, and DNA from the father was not available. This mutation leads to a frame-shift and a premature termination (PTC) codon at position 664 (L604Sfs*60 : NP_001167538). This insertion is not known in gnomAD. The gene is intolerant for loss of function (LoF) mutations, evidenced by a pLI value of 0,9. The mutation is localized in the αE helix in the FGFR1 tyrosine kinase domain, essential portion for its catalytic activity. Such a deletion should lead to loss-of-function. Alternatively, this mutation may lead to nonsense mediated mRNA decay (NMD), as predicted by Mutation taster and frequent for PTCs, causing haploinsufficiency.

Patient 2
The index case is the first child of two, of an unrelated Caucasian couple with no familial history of cleft, born after an uneventful pregnancy. Prenatal diagnosis of right cleft of the lip and alveolus was confirmed at birth. He was born at 39 WG, apgar score: 2/7, with following birth parameters: Weight: 3,680Kg, Height: 53 cm, Head Circumference: 34,5cm. Neonatal cranial, renal and cardiac ultrasounds were normal. Diagnosis of a submucous cleft palate was secondarily made. Apart from frequent upper airway infections, clinical examination at 2 years of age noted normal growth and development, and confirmed the isolated nature of his cleft. Parents' clinical examination was unremarkable.
A heterozygous nucleotide change (c.832C>T, NM_004087) in exon 10 of the *DLG1* gene was identified in the proband, and segregation analysis showed that the mother was a carrier. This nonsense variation was predicted to cause premature termination of protein synthesis at codon 278 (p.Arg278*, NP_004078) in the first of the three PDZ domains. The variant is not present in gnomAD. The gene is intolerant for LoF mutations as evidenced by a pLI value of 1.0. By deleting most of *DLG1* functional domains -PDZ domains, a src-homology 3 (SH3) domain, a band 4.1 domain, and a guanylate kinase homolog (GUK) domain- this mutation likely represents loss-of-function. Alternatively, this mutation may lead to NMD, as predicted by Mutation taster (Probability: 1) still causing haploinsufficiency.

Patient 3
The index case is the younger child of the two of an unrelated Caucasian couple. He was born at term. He had a left-sided microform of cleft lip, described as congenital healed cleft lip and an alveolar notch, associated with a cleft of the soft

and hard palate. His father has an isolated complete left-sided cleft affecting the lip, alveolus and palate. His paternal grand-father has an isolated cleft of the lip. Genetic tests included karyotype, 22q11.2 FISH, and *IRF6* sequencing and all were negative. Clinical examination at 6 years of age revealed normal growth and development.

We identified a heterozygous variation (c.1112delA, NM_004815) in exon 11 of *ARGHAP29* gene in both the proband and his father. Segregation study showed that his unaffected brother was a carrier and his affected paternal grand-father was not, being a phenocopy. This nonsense variation was predicted to cause premature termination of protein synthesis (p.Lys371fs, NP_004815) and it is not present in gnomAD. *ARGHAP29* is intolerant for LoF mutations as evidenced by a pLI value of 1.0. The mutation is located in the coiled coil region. Deletion of this part of the protein and the zinc finger domain likely causes loss of function. Alternatively this mutation may lead to NMD, as predicted by Mutation taster (Probability: 1), causing haploinsufficiency.

DISCUSSION

Phenotypic classification of CLP into 3 subgroups, defined as CL, CL with CP, and CP, fails to capture discontinuous CLP because "complete" CLP and "discontinuous" CLP cannot be distinguished. As discontinuous CLP is likely included in either CL or CL/P group, non-syndromic CLP genes have not been specifically implicated in discontinuous CLP. This WES study is the first dedicated to non syndromic discontinuous CLP. Deleterious mutations were found in 3 out of 8 patients, implicating *FGFR1*, *ARGHAP29* and *DLG1* genes. This represents almost 40% of this small cohort.

FGF signaling plays an essential role in palatal development. Missense and nonsense mutations in the *FGF* gene family, including *FGFR1*, may account as much as 3-5% of nsCLP cases [70]. Loss-of-function (LoF) mutations in *FGFR1* cause familial and sporadic Kallman syndrome (KS) and normosmic idiopathic hypogonadotropic hypogonadism. The main KS features are anosmia and hypogonadotrophic hypogonadism, and 5 to 10% of these patients have a cleft [19, 204]. Missense and nonsense *FGFR1* mutations have also been identified in nsCLP patients [70]. For example the p.R609X non-sense mutation cosegregated in a family with either KS and CLP or isolated CLP [70]. Patient 1 has non syndromic CLP. Following the discovery of this *FGFR1* mutation, a clinical evaluation

focusing on potential hypogonadotrophic hypogonadism and hypo/anosmia has been proposed.

Meta-analysis of GWAS data and a replication study identified 3q29 chromosomal locus as a cleft-susceptibility locus, with *DLG1* as a candidate gene for nsCL/P [205]. *DLG1* encodes a membrane-associated guanylate kinase protein (MAGUK), scaffolding protein that is involved in apical-basal cell polarity, cell-cell adhesion and regulation of cellular proliferation. *Dlg1* null mutant mice die soon after birth and exhibit major congenital birth defects with craniofacial dysmorphogenesis including cleft palate [206]. *Dlg1* is expressed in mesenchymal and epithelial cells throughout palate, and protein-protein interactions involving the C-terminal domains seem to be essential for the normal function within craniofacial and palatal morphogenesis. It is suggested that loss of Dlg1 may affect craniofacial development through Wnt/PCP signalling [207]. Individuals with a 3q29 interstitial microdeletion syndrome, including *DLG1* gene, have orofacial cleft in 4 to 9% [208]. Moreover, an oral cleft was reported in a patient with a frame shift mutation (c.338_339insTATGTAATTCATAACAATAATAAAT, NM_001098424, p.(E113Dfs*3) in Geno2MP (see URL). The mutation in patient 2 is associated with incomplete penetrance, as his mother is an unaffected carrier.

Arhgap29 encodes Rho GTPase activating protein 29, and is expressed in the palatal shelves and oral epithelium during craniofacial development. It was concluded that *ARHGAP29* is the etiologic gene for nsCL/P in the 1p22 locus [179]. Further sequencing of *ARHGAP29* in nsCL/P patients reported rare variants [179, 209-211] with loss-of-function (LoF) contributing to the pathology [211]. Exome sequencing conducted in an isolated CP family implicated one rare missense variant, with an in vivo assay in favour of LoF [212]. *ARHGAP29* variants are associated with an intra- and interfamilial variable expressivity, and in described families, follow an autosomal dominant inheritance pattern with incomplete penetrance, conferring a moderate risk to the disease. Involvement of the lip can be subtle, presenting as “healed” cleft lip in the index case of our family 3, and as a “congenital lip scar” in one of the familial cases previously reported [211]. Grandfather is a phenocopy. The family pedigree of patient 3 witnesses intrafamilial variable expressivity, demonstrating that discontinuous CLP, CL and CL/P are part of the same phenotypic spectrum.

In the sole reported large cohort of 356 syndromic and non-syndromic discontinuous CLP, four patients were clinically diagnosed as having Van der Woude syndrome (VWS) and one as having popliteal pterygium syndrome

(PPS)[202]. Since 70 % of VWS syndrome and 97% of PPS have an *IRF6* mutation, and taking into account that *IRF6* has been implicated in nsCL/P and nsCP, *IRF6* is a good candidate gene for discontinuous cleft. Interestingly, *irf6* was shown to regulate *Arhgap29* [179]. Yet, a direct targeted *IRF6* gene sequencing study of 4 individuals with discontinuous cleft was negative [203]. We did not find any *IRF6* mutation either.

Considering discontinuous CLP as a distinct group would have defined the 2 cases with familial history (table 1) as "mixed" families. Patient 3 had a discontinuous CLP with a microform of CL, a phenotype that could represent 25% (92/356) of discontinuous CLPs [202]. However, in discontinuous cleft, clefting of the secondary palate is not lateralized as in CL with cleft palate, but strictly median. Interestingly, individuals with oral orbicularis muscle (OOM) discontinuities, considered as a subclinical microform of CL, have been described in patients with CP [213]. They also likely belong to the discontinuous CLP group. As discontinuous CLP shows considerable variability in terms of affected anatomical structures (Table 1), an even more detailed subgrouping needs to be considered to enable more detailed study of this atypical phenotype.

Facial development is a complex multistep process that requires a fine spatio-temporal coordination of expression of numerous genes. Embryonic development of the primary and secondary palate can be subdivided into an early (4 to 7 weeks of gestation) and late period (7 to 12 weeks of gestation), with different exposures to genes and environmental factors. Formation of the upper lip is completed at 48 days of gestation, whereas palatal fusion is completed only by 12 weeks of gestation, with posterior palatal part being the last to form. According to embryological origins, the palate is divided into the anterior intermaxillary segment, and posterior or secondary palate. The anterior palate which extends posteriorly to the incisive foramen, or clinically to the incisive papilla, derived from the medial nasal processes, fuses posteriorly with the secondary palate derived from the maxillary processes [4]. Based on epidemiological data and differences in embryological origins, CL/P and CP have historically been considered as distinct entities. Nevertheless CL/P and CP can segregate in the same family, with etiologic mutations in genes such as *p63, MSX1, IRF6, and FGFR1* [13], and epidemiologic data suggest that cleft lip only may have unique etiologic features [62]. Moreover, the example of *Bmp1a* preferential expression in the primary palate and anterior secondary palate during palatal outgrowth [12] reflects similarities of spatio-temporal restricted expression pattern although different

embryological origins. Interestingly, *Grem1* gene, which encodes a secreted antagonist of the *Bmp4,* is expressed in the developing lip and soft palate, but not in the hard palate. Variations in the non-coding region near the *GREM1* gene show a highly significant association with discontinuous cleft [214]. *S*pecific mechanisms underlying the occurrence of discontinuous cleft have yet to be explained.

To date only few WES studies have been performed to search for rare coding variants in nsCLP individuals. Moreover, they were mostly applied to familial cases, in which genetic factors are more likely to be prevalent [183, 201, 212, 215-218]. Exome sequencing of 46 nsCLP families unraveled mutations in genes mutated in syndromic forms of CLP in 5 families (in 10%), including nsCP families [183]. In the current study, selection of individuals was exclusively made on the rare discontinuous CLP phenotype, whether familial or not. The mutation pick-up rate was high, as LoF mutations were found in 3 genes known to be implicated in nsCLP, including two in sporadic cases. WES study of homogeneously phenotyped individuals has been one of the keys for identifying genes in rare Mendelian disorders, including syndromic CLP genes [82, 83]. Our findings call for accurate phenotyping and stratification in complex disorders, such as nsCLP, and a meticulous clinical evaluation has to be proposed prior to genetic studies.

Empiric recurrence risk and gender ratio are different between non syndromic CL/P and CP. Substantial evidence is accumulating on rare variants contributing to nsCLP, and finding a mutation in a Mendelian CLP-disorder can consistently enhance the recurrence risk of nsCLP. Familial studies have proven to be an effective strategy to identify rare variants and medical screening of such patients has been proposed [183, 212, 219]. Similarly, considering the results on this study, screening patients with discontinuous cleft should be taken into consideration in genetic counseling. However, in familial cases rare variants and nsCLP often cosegregate in an autosomal dominant manner with incomplete penetrance. For heterozygous LoF variants in *ARHGAP29,* penetrance was calculated at 59%, indicating a moderate risk to develop nsCLP [211]. Such data need to be collected to estimate recurrence risk for each gene, essential for personalized counseling.

CONCLUDING REMARKS

Whole exome sequencing of clinically well-defined subgroups of CLP is a relevant approach to study CLP etiopathogenesis. It could help more accurate clinical,

epidemiological and fundamental research, ultimately resulting in better diagnosis and care of CLP patients.

Acknowledgment

We are grateful to all the family members for their invaluable contributions. These studies were partially supported by funding from the Belgian Science Policy Office Interuniversity Attraction Poles (BELSPO-IAP) program through the project IAP P7/43-BeMGI; BridgeIris RBC/2013-PFS-EH-11; the Fonds de la Recherche Scientifique - FNRS; CdR: J.0080.16) (all to M.V.). We also acknowledge the support of la Communauté française de Wallonie-Bruxelles, la Lotterie Nationale, Belgium and la région Hauts de France, France. The authors thank the Genomics Platform of Université catholique de Louvain for Next Generation Sequencing and the Foundation against Cancer, Belgium.

TABLE LEGENDS:

Table 1: Summary of cases

Web resources

Highlander: http://sites.uclouvain.be/highlander/
Online Mendelian Inheritance in Man (OMIM): http://www.omim.org
1000 Genomes: http://www.1000genomes.org/
Genome of the Netherlands (GoNL): http://www.nlgenome.nl/
NHLBI Exome Sequencing Project (ESP): http://evs.gs.washington.edu/EVS/
Exome Aggregation Consortium (ExAC): http://exac.broadinstitute.org/
Genome Aggregation Database (gnomAD): http://gnomad.broadinstitute.org/
The Human Gene Mutation Database (HMGD): http://www.hgmd.cf.ac.uk/
Polyphen2: http://genetics.bwh.harvard.edu/pph2/
SIFT: http://sift.jcvi.org/
Mutation Taster: http://www.mutationtaster.org/
Mutation Assessor: http://mutationassessor.org/
Functional Analysis though Hidden Markov Models (FATHMM): http://fathmm.biocompute.org.uk/
denovo-db:http://denovo-db.gs.washington.edu/denovo-db/
Mutalyzer: https://mutalyzer.nl/
Genotype to Mendelian Phenotype (Geno2MP):https://geno2mp.gs.washington.edu

Bibliography

1. Dixon, M.J., et al., *Cleft lip and palate: understanding genetic and environmental influences.* Nat Rev Genet, 2011. **12**(3): p. 167-78.
2. Ng, S.B., et al., *Exome sequencing identifies the cause of a mendelian disorder.* Nat Genet, 2010. **42**(1): p. 30-5.
3. Gilissen, C., et al., *Unlocking Mendelian disease using exome sequencing.* Genome Biol, 2011. **12**(9): p. 228.
4. Bureau, A., et al., *Whole exome sequencing of distant relatives in multiplex families implicates rare variants in candidate genes for oral clefts.* Genetics, 2014. **197**(3): p. 1039-44.
5. Basha, M., et al., *Whole exome sequencing identifies mutations in 10% of patients with familial non-syndromic cleft lip and/or palate in genes mutated in well-known syndromes.* J Med Genet, 2018.
6. Neiswanger, K., et al., *Whorl patterns on the lower lip are associated with nonsyndromic cleft lip with or without cleft palate.* Am J Med Genet A, 2009. **149A**(12): p. 2673-9.
7. Jiang, R., J.O. Bush, and A.C. Lidral, *Development of the upper lip: morphogenetic and molecular mechanisms.* Dev Dyn, 2006. **235**(5): p. 1152-66.
8. Alvarez, C.W., M.L. Guion-Almeida, and A. Richieri-Costa, *Clinical and genetic study on 356 Brazilian patients with a distinct phenotype of cleft lip and palate without alveolar ridge involvement.* J Craniomaxillofac Surg, 2014. **42**(8): p. 1952-7.
9. Cuddapah, S.R., et al., *IRF6 Sequencing in Interrupted Clefting.* Cleft Palate Craniofac J, 2016. **53**(3): p. 373-6.
10. Jugessur, A., et al., *Genetic determinants of facial clefting: analysis of 357 candidate genes using two national cleft studies from Scandinavia.* PLoS One, 2009. **4**(4): p. e5385.
11. Riley, B.M., et al., *Impaired FGF signaling contributes to cleft lip and palate.* Proc Natl Acad Sci U S A, 2007. **104**(11): p. 4512-7.
12. Dode, C., et al., *Loss-of-function mutations in FGFR1 cause autosomal dominant Kallmann syndrome.* Nat Genet, 2003. **33**(4): p. 463-5.
13. Pitteloud, N., et al., *Mutations in fibroblast growth factor receptor 1 cause both Kallmann syndrome and normosmic idiopathic hypogonadotropic hypogonadism.* Proc Natl Acad Sci U S A, 2006. **103**(16): p. 6281-6.
14. Mostowska, A., et al., *Common variants in DLG1 locus are associated with non-syndromic cleft lip with or without cleft palate.* Clin Genet, 2018. **93**(4): p. 784-793.
15. Caruana, G. and A. Bernstein, *Craniofacial dysmorphogenesis including cleft palate in mice with an insertional mutation in the discs large gene.* Mol Cell Biol, 2001. **21**(5): p. 1475-83.

16. Rivera, C., et al., *Requirement for Dlgh-1 in planar cell polarity and skeletogenesis during vertebrate development.* PLoS One, 2013. **8**(1): p. e54410.
17. Willatt, L., et al., *3q29 microdeletion syndrome: clinical and molecular characterization of a new syndrome.* Am J Hum Genet, 2005. **77**(1): p. 154-60.
18. Leslie, E.J., et al., *Expression and mutation analyses implicate ARHGAP29 as the etiologic gene for the cleft lip with or without cleft palate locus identified by genome-wide association on chromosome 1p22.* Birth Defects Res A Clin Mol Teratol, 2012. **94**(11): p. 934-42.
19. Chandrasekharan, D. and A. Ramanathan, *Identification of a novel heterozygous truncation mutation in exon 1 of ARHGAP29 in an Indian subject with nonsyndromic cleft lip with cleft palate.* Eur J Dent, 2014. **8**(4): p. 528-32.
20. Leslie, E.J., et al., *Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci.* Am J Hum Genet, 2015. **96**(3): p. 397-411.
21. Savastano, C.P., et al., *Impact of rare variants in ARHGAP29 to the etiology of oral clefts: role of loss-of-function vs missense variants.* Clin Genet, 2017. **91**(5): p. 683-689.
22. Liu, H., et al., *Exome sequencing provides additional evidence for the involvement of ARHGAP29 in Mendelian orofacial clefting and extends the phenotypic spectrum to isolated cleft palate.* Birth Defects Res, 2017. **109**(1): p. 27-37.
23. Weinberg, S.M., et al., *Rethinking isolated cleft palate: evidence of occult lip defects in a subset of cases.* Am J Med Genet A, 2008. **146A**(13): p. 1670-5.
24. Grosen, D., et al., *A cohort study of recurrence patterns among more than 54,000 relatives of oral cleft cases in Denmark: support for the multifactorial threshold model of inheritance.* J Med Genet, 2010. **47**(3): p. 162-8.
25. Baek, J.A., et al., *Bmpr1a signaling plays critical roles in palatal shelf growth and palatal bone formation.* Dev Biol, 2011. **350**(2): p. 520-31.
26. Ludwig, K.U., et al., *Meta-analysis Reveals Genome-Wide Significance at 15q13 for Nonsyndromic Clefting of Both the Lip and the Palate, and Functional Analyses Implicate GREM1 As a Plausible Causative Gene.* PLoS Genet, 2016. **12**(3): p. e1005914.
27. Pengelly, R.J., et al., *Deleterious coding variants in multi-case families with non-syndromic cleft lip and/or palate phenotypes.* Sci Rep, 2016. **6**: p. 30457.
28. Aylward, A., et al., *Using Whole Exome Sequencing to Identify Candidate Genes With Rare Variants In Nonsyndromic Cleft Lip and Palate.* Genet Epidemiol, 2016. **40**(5): p. 432-41.

29. Hoebel, A.K., et al., *Candidate Genes for Nonsyndromic Cleft Palate Detected by Exome Sequencing*. J Dent Res, 2017. **96**(11): p. 1314-1321.
30. Holzinger, E.R., et al., *Analysis of sequence data to identify potential risk variants for oral clefts in multiplex families.* Mol Genet Genomic Med, 2017. **5**(5): p. 570-579.
31. Ng, S.B., et al., *Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome.* Nat Genet, 2010. **42**(9): p. 790-3.
32. Leslie, E.J., et al., *IRF6 mutation screening in non-syndromic orofacial clefting: analysis of 1521 families.* Clin Genet, 2016. **90**(1): p. 28-34.

Patient	Gender	Familial history of CLP	Discontinuous CLP description		Mutation
			cleft of the primary palate	cleft of the secondary palate	
1	F	-	right labioalveolar	submucous	*FGFR1*, c.1809_1810insTC
2	M	-	left labioalveolar	submucous	*DLG1*, c.832C>T
3	M	+	right intra-uterine healed lip and alveolar	soft and hard palate	*ARHGAP29*, c.1112delA
4	M	-	right labioalveolar	submucous	
5	M	-	left labioalveolar	submucous	
6	M	-	right labial	soft and hard palate	
7	M	-	left labial	soft and hard palate	
8	M	+	right labial	submucous	

Table 1: Summary of cases

D. Introductory remarks on preliminary results

"LRRC1 gene is mutated in Van der Woude syndrome and isolated CP"

Bénédicte Demeer, Nicole Revencu, Raphael Helaers, Cica Gbaguidi, Geneviève Francois, Bernard Devauchelle, Bénédicte Bayet, Miikka Vikkula

in preparation

I selected the patients and performed the WES data analysis, the validation/co-segregation of identified variants, the RNA and the protein studies. Together with NR, I clinically reassessed patients and families, and collected blood samples required for further assays.
Sequencing of 15 VWS patients (negative for *IRF6*) is ongoing. We plan to collaborate to use animal models such as zebrafish to study the consequence of mutated *LRRC1* on orofacial development.

LRRC1 GENE IS MUTATED IN VAN DER WOUDE SYNDROME AND ISOLATED CP

Bénédicte Demeer[1,2,3], Nicole Revencu[1,4], Raphael Helaers[1], Cica Gbaguidi [5], Geneviève Francois[6], Bernard Devauchelle[3,5], Bénédicte Bayet[7], Miikka Vikkula[1]

[1] Human Molecular Genetics, de Duve Institute, University of Louvain, Brussels, Belgium

[2] Center for Human Genetics, CLAD nord de France, CHU Amiens, Amiens, France

[3] EA CHIMERE, Université Picardie Jules Verne, Amiens, France

[4] Center for Human Genetics, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

[5] Department of Maxillofacial Surgery and Stomatology, centre de compétence fentes et malformations faciales (MAFACE), CHU Amiens-Picardie, Amiens, France

[6] Department of Pediatrics, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

[7] Centre Labiopalatin, Division of Plastic Surgery, Cliniques universitaires Saint-Luc, University of Louvain, Brussels, Belgium

Correspondence to
Professor Miikka Vikkula,
Human Molecular Genetics
de Duve Institute, University of Louvain
Brussels, Belgium
e-mail: miikka.vikkula@uclouvain.be
phone: + 32-2-764 7490 / fax: + 32-2-764 7460

KEYWORDS: *LRRC1,* Van der Woude, cleft palate, WES, *DLG1*, cleft lip and/or palate

Word count: 2435

ABSTRACT:

Van der Woude syndrome (VWS) , the most common cleft syndrome is characterizd by cleft of the lip and/or palate (CLP) and lower lip pits. The clinical presentation of VWS is variable and can present as an isolated CLP. Approximately 75 % of VWS are explained by mutation in one of the causative genes : *IRF6* or *GRHL3* leaving 25% cases of VWS unexplained. We conducted Whole Exome sequencing (WES) on 16 non-syndromic CLP families and one VWS family in order to identify rare genetic causative variants. We found one novel gene, LRRC1, in common for a non-syndromic cleft palate family and the VWS family. The two LRRC1 nonsense mutations identified result in a mutant truncated stable protein missing the C terminal PDZ binding site of the protein.

INTRODUCTION

Cleft of the lip and/or palate (CLP) are among the most common birth defects, with an approximate incidence of 1/700 live births and with a wide variability in expressivity depending on ethnicity, gender and cleft type. Although not associated with a consequent rate of mortality in developed countries, they represent a significant lifelong morbidity. This requires multidisciplinary treatment and care [13].

~30 % of CLPs occur in syndromes, which associate additional developmental, structural and/or cognitive anomalies and generally follow Mendelian inheritance with an important variability in expressivity and incomplete penetrance. The most common cleft syndrome is Van der Woude syndrome (VWS) (MIM 119300). Mutations in interferon regulatory factor 6 (*IRF6*, MIM 607199) account for ~70% of individuals with VWS, and another 5% are found to be mutated in grainy head-like 3 (*GRHL3)* gene (MIM 608317). The remaining 25% cases of VWS have no known causal mutation. Incomplete penetrance is observed in Van der Woude syndrome and VWS is phenotypically distinguished from nsCLP only by "minor" clinical signs such as pits of the lower lip or hypodontia, which are also variably present in VWS patients: 20% do not have lip pits. Thus, 10 to 15 % of VWS appears as nsCLP [219].

Given the complex and heterogeneous nature of non-syndromic clefts, different genetic approaches have been employed to identify loci and genes, including: linkage analysis in families, association studies using case/parent trios or case-control cohorts for candidate-loci or for genome-wide association studies, identification of chromosomal abnormalities or microdeletions in individual cases, and sequencing of candidate genes in affected individuals. They have all had successful applications [13]. Generation and analyses of animal models has also contributed to the understanding of underlying causes of non-syndromic CLP.

With rapid improvement in technology, in particular next-generation sequencing and bioinformatic algorithms, we now have the ability to interrogate the entire exome or genome in order to identify risk loci. Whole exome sequencing (WES) has been successfully performed for Mendelian disorders, becoming a useful approach for proving clinical molecular diagnosis [200]. This strategy has also been applied to identify causes of complex traits, including cleft lip and palate [183, 201]. We used WES to identify rare genetic causative variants in a cohort of nsCLP families and one VWS family, and identified one novel gene in common for 2 families.

MATERIALS AND METHODS

PATIENT SELECTION

Clinical data and samples from patients with nsCLP and their family members were collected in collaboration with the Centre labiopalatin, Cliniques universitaires St Luc, Brussels, Belgium and Amiens-Picardie Hospital, France. Informed consent was signed by each participant prior to conducting the study, as approved by the institutional review boards. Each patient was asked to fill in a questionnaire recapitulating personal, familial history and his/her malformations. The CLP phenotyping was based on clinical examination by the plastic surgeon while evaluating the infant before surgery and follow up, and on the questionnaire. A total of 30 affected patients were included from 16 ns CLP families, with index patients only for 7 families, and 3 affected for 1 family, and from 1 Van der Woude family Affected subjects were diagnosed as having : cleft lip : 4, cleft lip with cleft palate: 14, cleft palate: 10, Van der Woude syndrome (VWS): 2 (same family). All patients had a normal standard karyotype and many had been tested negative for 22q11.2 deletion and *IRF6* mutation. NsCP families had also been tested for a rare recurrent mutation in *GRHL3*: c.1171C>T (NM_198174.2); p.Arg391Cyst (NP_002327.2)

WHOLE EXOME SEQUENCING (WES)

DNA was extracted from blood samples using Wizard® Genomic DNA Purification Kit (Promega). WES was carried out by using 1 µg of genomic DNA per sample. Exomes were captured using a commercial enrichment kit (Agilent SureSelectXT Human All Exon KitV5) and enriched libraries were sequenced on an Illumina Hiseq2000. Generated reads were aligned to the human genome build hg19 using Burrows-Wheeler Aligner (BWA 0.7.15; see URLs). Data processing and variant calling were performed with Picard 1.107, Genome Analysis Toolkit 3.3 package (GATK) respectively. The generated variant (.vcf) files were imported into a database and further analysed using Highlander 14.10.3, an in-house bioinformatic framework for variant annotation, filtering, and visualization (see URLSs) (Helaers et al., under revision).

We considered as rare variants those with a frequency below 0.5% in public databases (ExAC, 1000 Genomes and GO-NL (genome of the Netherlands) and in our in-house control database. As “likely pathogenic” variants, we considered those with a high impact (canonical splice-site variants and those generating a premature termination codon=PTC (out-of-frame indels and nonsense variants)), or a

moderate impact (non synonymous-NS), as annotated by SnpEff 4.1 and a Combined Annotation dependent Depletion (CADD) phred score > 15. Missense variants were considered as possibly pathogenic only if predicted to be possibly/probably damaging in at least three out of seven in silico variant classifiers (Sift, Polyphen, LRT, Mutation taster, Mutation assessor, FATHMM, DEOGEN). We first considered a list of 361 genes previously implicated in any form of clefting as candidates [184]. We secondarily searched for genes with candidate variants in ≥ 2 families.
For all variants that satisfied our criteria, we performed co-segregation analysis on all the affected and non-affected family members that were available. Variants were PCR-amplified followed by sequencing on an ABI 3130XL genetic analyzer (Applied BioSystems).

VALIDATION OF mRNA STABILITY

RNA was extracted from blood samples (lymphocytes) with TriPure (Roche) and retrotranscribed using RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas) with random hexamers. PCR-amplification was done on cDNA using a *LRRC1* gene primer pair, forward in exon 12 (CCTGTGTCTTACTTCCTCAG) and reverse in exon 14 (TTACCTCCTGTGTCTTCCTC). Amplicon's size (378 bp) was evaluated with an agarose gel, using GeneRuler 100 bp (Thermo Scientific) DNA ladder

CELL CULTURE AND EXPRESSION OF RECOMBINANT LRRC1

COS-7 cells were maintained in DMEM supplemented with 10% FBS. The pCMV3-LRRC1 expression vectors containing full-length Myc-DDK-tagged human LRRC1 cDNA (RC200125, Origene Inc.) was used. Mutations were generated using the QuickChange Site-Directed Mutagenesis method (Agilent). Primers were designed using QuikChange Primer Design (Agilent). We separately introduced two identified *LRRC1* mutations and a deletion of the C-terminal TSV-motif, used as a control. Mutant strand synthesis was initiated with 10ng of vector DNA and addition of 3µL DMSO. Cycling was adjusted to 16 (30 sec. denaturation at 95°C, 1 min. annealing at 59°C, 9 minutes elongation at 68°C) and an extra 9 min. elongation at 67°C. The amplification products were transformed in JM109 competent bacteria and the entire LRRC1 cDNA sequence was verified by Sanger sequencing. Plasmids were transfected in 80-95% confluent COS-7 cells using JetPEI (Polyplus transfection) in 6-well plates. Cells were co-transfected with 3 µg

of one of the LRRC1-containing plasmids (WT, del TSV p.Arg459*, p.Leu486*) Six hours after transfection, the medium was changed to DMEM without phenol red (Thermofisher), 1% FBS and 1mM pyruvate, and cultured for a further 3 days.

IMMUNOBLOTTING

To analyze expression and secretion of recombinant LRRC1, conditioned media were collected after 3 days and cells were lysed during 30 min at 4°C with 50 mM Tris/HCl (pH 8.0), 150mM NaCl, 1% Nonidet P-40, 20mM EDTA supplemented with Complete, Mini protease inhibitor cocktail tablets (Roche) and PhosSTOP (Roche). Cell lysates were sonicated for 15 sec. and centrifuged at 15,000*g* for 20min. Conditioned media were centrifuged at 400*g* for 5min. and the pellet of cells was removed. Total protein concentrations were quantified using Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific). Samples were separated on SDS/PAGE, transferred onto PVDF membranes (Merck Millipore) and immunoblotted with a human LRRC1 antibody (LS-C392798, LSBio) or anti-Flag (F1804, Sigma Aldrich) or α-Actinin antibody (6487, cell signaling technology) and probed with secondary antibody (HRP conjugated anti-goat IgG, HRP conjugated anti-mouse IgG, or HRP conjugated anti-goat IgG respectively). Bands were detected with enhanced chemiluminescence reagent (Perkin Elmer) and Supersignal West Femto Maximum Sensitivity substrate (Thermo Fisher Scientific).

RESULTS:

CLINICAL PHENOTYPE

Family 1: VWS family

The proband is the second child of an unrelated Caucasian couple, with family history of Van der Woude syndrome: the father and one paternal aunt have an abnormal uvula, her uncle a cleft of the palate (CP) and lip pits, one great uncle a CP with one his daughters said to have a CP, and her great-great grand-father had a CP. Diagnosis of VWS, associating CP and lip pits was made at birth. She had neonatal feeding difficulties. A type I diabetes was diagnosed at the age of 5 years. At 15 years of age, she has a normal development and normal growth: height: 1,61m, weight 76kg. Her elder sister, who had a normal clinical and endobuccal examination, developed type II diabetes at the age of 15 years.

Family 2: nsCP

The index case is the elder of two, of an unrelated Caucasian couple with no known familial history, born after an uneventful pregnancy. She was diagnosed at birth with a cleft of the soft and hard palate. Her 3-years-younger sister was diagnosed of having cleft of the soft palate. They both have normal growth and development. Parents' clinical examination noted a maternal submucous cleft palate.

IDENTIFICATION OF LRRC1 MUTATIONS AND COSEGREGATION ANALYSIS

In family 1, we identified a c.1375C>T; p.Arg459* (NM_018214); (NP_060684.4) mutation in exon 13 of the *LRRC1* gene in the proband (IV-19) and her great uncle (II-2). Segregation study of the available DNAs showed that II-5, III-7, III-8, III-9 were carriers and II-6, III-6, III-10, IV-20 were not. This nonsense mutation is not known in gnomAD. The pLI value of LRRC1 in EXAC is 0.59 . The mutation may lead to a strongly truncated protein (-66 AA/more than 10% missing) lacking the coiled-coil domain and the TSV-motif. Alternatively this mutation might lead to nonsense mediated mRNA decay (NMR).

In family 2, we identified a c.1457T>G; p.Leu486*(NM_018214); (NP_060684.4) mutation in exon 14 of the *LRRC1* gene in the proband and her sister. Segregation study found that their mother was carrier. This nonsense mutation is not known in gnomAD. The mutation may lead to a slightly truncated protein (-39 AA/less than 10% missing) lacking part of the coiled-coil domain and the TSV-motif. Alternatively, this mutation might also lead to NMR, as known for many PTC-generating mutations.

RNA STUDY

We observed a unique amplicon obtained by RT-PCR performed in mRNA obtained from the patient's lymphocytes. Sequencing revealed equal presence of the wild type and mutated alleles (Figure 2)

PROTEIN STUDY

We detected by Western blot LRRC1 protein extracted from Cos7 cells transfected with the various constructs encoding mutated LRRC1: c.1375C>T (p.Arg459*) and c.1457T>G (p.Leu486*), and TSV-deleted variant. The truncating mutations led to production of shorter proteins but with quantities similar to that of the wild-type or the TSV-deleted forms (Figure 3).

DISCUSSION

Family 1 has Van der Woude syndrome with phenotype ranging from isolated minor involvement of the soft palate (abnormal uvula) (III-12, III-13) to cleft of the soft and hard palate with lip pits (I-1, II-2). Family 2 also shows a great range of intrafamilial variability in expression, ranging from submucous cleft palate (I-2) to cleft of the soft and hard palate (II-2). No clinical sign of Van der Woude syndrome was observed, but as 10 to 15 % of VWS mimic nsCLP, a VWS diagnosis should not be excluded considering the limited pedigree size and incomplete penetrance. Interestingly both families have clefts of the palate only.

LRRC1 is ubiquitously expressed in adult human tissues, but no data is available regarding embryogenesis. Data obtained in mouse embryos (Expression Atlas, see URL) show a medium-level expression in head mesenchyme, mandibular arch and maxillary arch at E9.5, and a low expression in lateral nasal prominence, medial nasal prominence, maxillary arch and mandibular arch, at E10.5, prior to palatogenesis. Orofacial expression profile is not documented after E10.5. Further expression analyses are thus needed.

LRRC1, or LANO (LAP and no PDZ), belongs to the LAP (Leucine rich repeats (LRR) and PDZ) protein family, and contains 16 leucine-rich repeats (LRRs), and a LAP-specific domain (LAPS) carboxy-terminal to the LRRs, but no PDZ domain. A general feature of LAP proteins is their basolateral localization mediated by the 16 LRRs [220]. LAPs are involved in establishing epithelial cell polarity and they act as adaptors in different intracellular signaling pathways [220, 221]. Saito and co-workers demonstrated that LRRC1 and more specifically, a C-terminal TSV-motif, directly interact with the PDZ domains of membrane-associated guanylate kinase (MAGUK) proteins, including DLG1 and PSD95 (DLG4) in epithelial cells. A second pool of LRRC1 is complexed with ERBIN. Deletion of the C-terminal TSV-motif abrogates LRRC1 interactions with PDZ domains [222]. The LRRC1 nonsense mutations c.1375C>T and c.1457T>G result in a mutant truncated stable protein with 16 LRRs but lacking the TSV-motif. These proteins would thus likely be located at the basolateral side of epithelial cells, but without being able to interact with PDZ domains.

LRRC1 appears to be a good candidate for clefts through its implication in various protein complexes. LRRC1 and DLG1 belong to the same protein complex [222]. *DLG1* is a candidate gene for nsCL/P [205]. *DLG1* encodes a scaffolding protein MAGUK, which is involved in apical-basal cell polarity, cell-cell adhesion and

regulation of cellular proliferation. *Dlg1* null mutant mice die soon after birth and exhibit major congenital birth defects with craniofacial dysmorphogenesis including cleft palate [206]. *Dlg1* is expressed in mesenchymal and epithelial cells throughout palate, and protein-protein interactions involving the C-terminal domains seem to be essential for the normal function during craniofacial and palatal morphogenesis. It is suggested that loss of Dlg1 may affect craniofacial development through Wnt/PCP signalling [207]. Individuals with a 3q29 interstitial microdeletion syndrome, including the *DLG1* gene have orofacial clefts in 4 to 9% [208]. Moreover, a patient with an oral cleft and a frame shift mutation in *DLG1* p.(E113Dfs*3) has been described (ref). We recently reported a nsCLP patient with a stop mutation in *DLG1* (Demeer et al., submitted).

LRRC1 indirectly interacts with ERBIN. ERBIN is a member of the LAP family. In its intermediate region between the N-terminal LRRs and the C-terminal PDZ-domains, it has a specific SID-domain (Smad-interacting domain or Smad-inhibitory domain), which is crucial for Erbin-R-Smad2/Erbin-R-Smad3 interaction, as well as for Erbin's inhibitory role in TGFB signaling [223]. Within the orofacial region, differentiation of the medial edge epithelial cells via epithelia-mesenchymal transformation occurs subsequent to the fusion of the secondary palatal shelves and it has been shown to be regulated by TGFB [224]. The role of Erbin in TGFB-mediated palatal fusion and epithelial transformation is unknown.

CONCLUDING REMARKS

The present study suggests *LRRC1* as a candidate gene for both Van der Woude syndrome and non-syndromic cleft palate. Additional cohort studies as well as expression pattern profiling in embryonic orofacial development and functional assays are needed to provide insight into the potential role of *LRRC1* in craniofacial development.

Acknowledgment

We are grateful to all the family members for their invaluable contributions. These studies were partially supported by funding from the Belgian Science Policy Office Interuniversity Attraction Poles (BELSPO-IAP) program through the project IAP P7/43-BeMGI; BridgeIris RBC/2013-PFS-EH-11; the Fonds de la Recherche Scientifique - FNRS; CdR: J.0080.16) (all to M.V.). We also acknowledge the support of la Communauté française de Wallonie-Bruxelles, la Lotterie Nationale, Belgium and la région Hauts de

France, France. The authors thank the Genomics Platform of Université catholique de Louvain for Next Generation Sequencing and the Foundation against Cancer, Belgium.

TABLE LEGENDS:

Figure 1. Pedigree and segregation of the *LRRC1* mutation.
A) family 1.
B) family 2
Figure 2. Sanger sequencing on LRRC1 cDNA
Figure 3. LRRC1 expression in cells expressing wild type or mutated LRRC1

Web resources

Highlander: http://sites.uclouvain.be/highlander/
Online Mendelian Inheritance in Man (OMIM): http://www.omim.org
1000 Genomes: http://www.1000genomes.org/
Genome of the Netherlands (GoNL): http://www.nlgenome.nl/
NHLBI Exome Sequencing Project (ESP): http://evs.gs.washington.edu/EVS/
Exome Aggregation Consortium (ExAC): http://exac.broadinstitute.org/
Genome Aggregation Database (gnomAD): http://gnomad.broadinstitute.org/
The Human Gene Mutation Database (HMGD): http://www.hgmd.cf.ac.uk/
Polyphen2: http://genetics.bwh.harvard.edu/pph2/
SIFT: http://sift.jcvi.org/
Mutation Taster: http://www.mutationtaster.org/
Mutation Assessor: http://mutationassessor.org/
Functional Analysis though Hidden Markov Models (FATHMM): http://fathmm.biocompute.org.uk/
denovo-db: http://denovo-db.gs.washington.edu/denovo-db/
Mutalyzer: https://mutalyzer.nl/
Expression atlas: https://www.ebi.ac.uk

Bibliography

1. Dixon, M.J., et al., *Cleft lip and palate: understanding genetic and environmental influences.* Nat Rev Genet, 2011. **12**(3): p. 167-78.
2. Leslie, E.J., et al., *IRF6 mutation screening in non-syndromic orofacial clefting: analysis of 1521 families.* Clin Genet, 2016. **90**(1): p. 28-34.
3. Gilissen, C., et al., *Unlocking Mendelian disease using exome sequencing.* Genome Biol, 2011. **12**(9): p. 228.
4. Bureau, A., et al., *Whole exome sequencing of distant relatives in multiplex families implicates rare variants in candidate genes for oral clefts.* Genetics, 2014. **197**(3): p. 1039-44.
5. Basha, M., et al., *Whole exome sequencing identifies mutations in 10% of patients with familial non-syndromic cleft lip and/or palate in genes mutated in well-known syndromes.* J Med Genet, 2018.
6. Jugessur, A., et al., *Genetic determinants of facial clefting: analysis of 357 candidate genes using two national cleft studies from Scandinavia.* PLoS One, 2009. **4**(4): p. e5385.
7. Legouis, R., et al., *Basolateral targeting by leucine-rich repeat domains in epithelial cells.* EMBO Rep, 2003. **4**(11): p. 1096-102.
8. Kravic, B., et al., *LAP proteins are localized at the post-synaptic membrane of neuromuscular junctions and appear to modulate synaptic morphology and transmission.* J Neurochem, 2016. **139**(3): p. 381-395.
9. Saito, H., et al., *Lano, a novel LAP protein directly connected to MAGUK proteins in epithelial cells.* J Biol Chem, 2001. **276**(34): p. 32051-5.
10. Mostowska, A., et al., *Common variants in DLG1 locus are associated with non-syndromic cleft lip with or without cleft palate.* Clin Genet, 2018. **93**(4): p. 784-793.
11. Caruana, G. and A. Bernstein, *Craniofacial dysmorphogenesis including cleft palate in mice with an insertional mutation in the discs large gene.* Mol Cell Biol, 2001. **21**(5): p. 1475-83.
12. Rivera, C., et al., *Requirement for Dlgh-1 in planar cell polarity and skeletogenesis during vertebrate development.* PLoS One, 2013. **8**(1): p. e54410.
13. Willatt, L., et al., *3q29 microdeletion syndrome: clinical and molecular characterization of a new syndrome.* Am J Hum Genet, 2005. **77**(1): p. 154-60.
14. Dai, F., et al., *Erbin inhibits transforming growth factor beta signaling through a novel Smad-interacting domain.* Mol Cell Biol, 2007. **27**(17): p. 6183-94.
15. Warner, D.R., et al., *Identification of three novel Smad binding proteins involved in cell polarity.* FEBS Lett, 2003. **539**(1-3): p. 167-73.

Figure 1. Pedigree and segregation of the *LRRC1* mutation

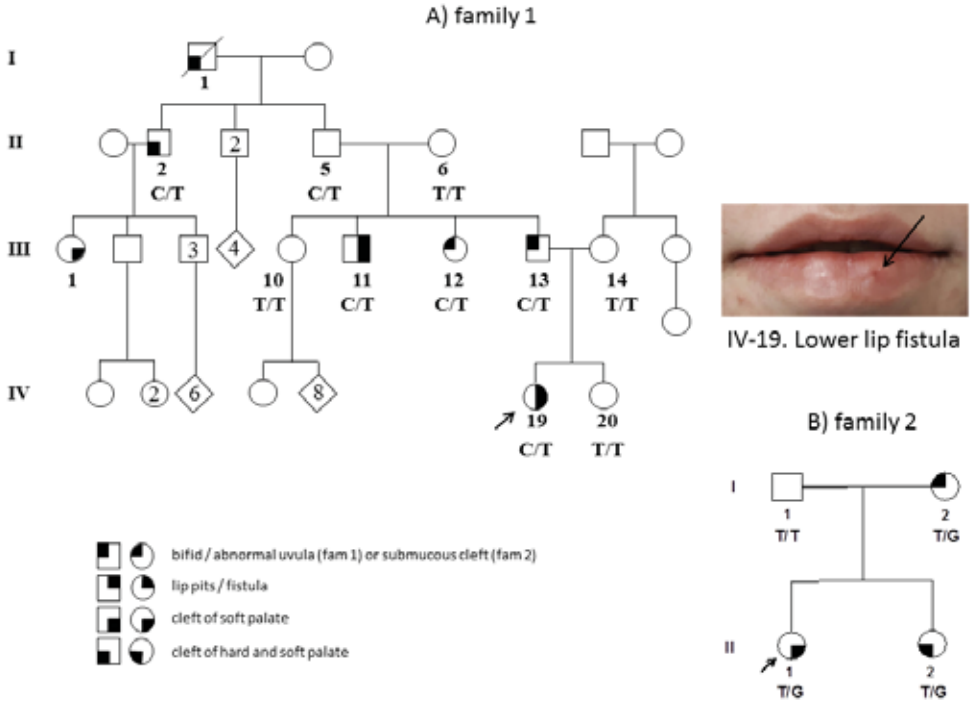
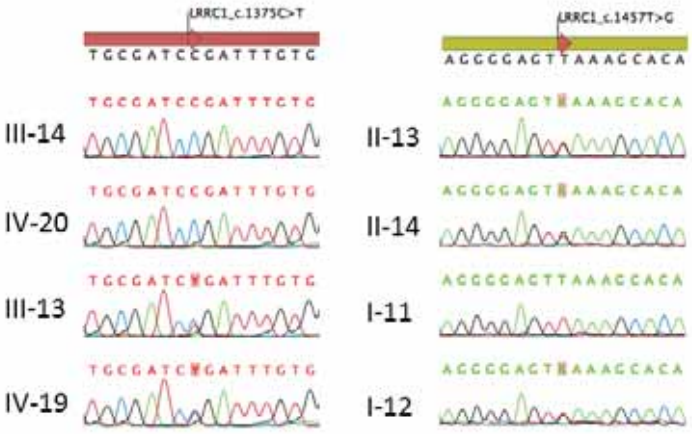


Figure 2. Sanger sequencing on LRRC1 cDNA



A) Family 1. Sequencing chromatograms: equal presence of wild type and mutated allele in individuals III-13 and IV-19

B) Family 2. Sequencing chromatograms: equal presence of wild type and mutated allele in individuals II-13, II-14 and I-2

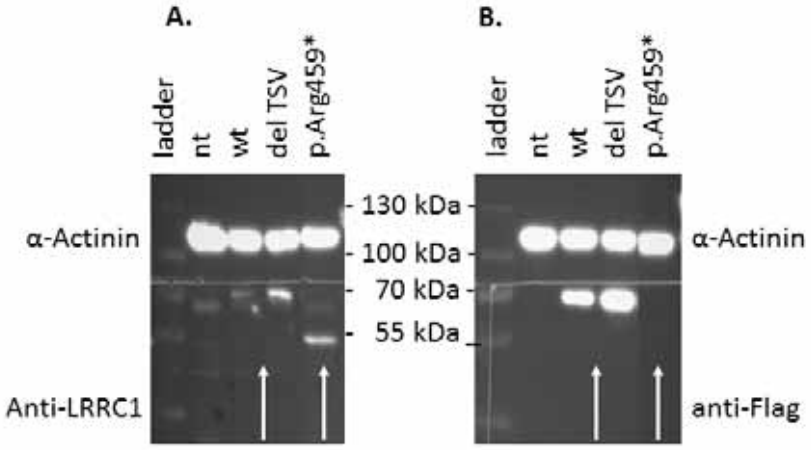


Figure 3. LRRC1 expression in cells expressing wild type or mutated LRRC1 (del TSV or p.Arg459*). A.) LRRC1 detected with the anti-LRRC1 antibody at around 59 kDa (arrow) for wt and del TSV, and at 51 kDa for p.Arg459*. Quantity similar to that of the wild-type or TSV-deleted form. NT: non-transfected. B.) LRRC1 detected with the anti-FLAG antibody at around 64 kDa (arrow) for wt and del TSV, but not for p.Arg459*.

V. DISCUSSION

A. Genomic era and nsCLP

The etiology of oral clefts is complex and implies both major and minor genetic influences with variable interactions from environmental factors. Segregation analysis indicated that the complex familial pattern observed in nsCLP is best explained as a mixture of monogenic cases probably dominantly inherited, combined with others which have a multifactorial etiology [66].

1. Detecting genomic alterations in nsCLP patients

Given the complex and heterogeneous nature of non-syndromic cleft, different genetic approaches have been employed to identify loci and genes, including: linkage analysis in families, association studies using case/parent trios or case-control cohorts for candidate-loci or for genome-wide association studies, identification of chromosomal abnormalities or microdeletions in individual cases, and sequencing of candidate genes in affected individuals. They all had successful applications [13]. However, it is estimated that identified regions and loci only account for 20–25% of the heritability, implying that additional variation should be identifiable through sequencing and other approaches. The development of next generation technology gives us the unique opportunity to interrogate concurrently the entire exome (whole exome sequencing, WES) and genome (whole genome sequencing, WGS) so as to identify risk loci. WES has been successfully performed for Mendelian disorders [83].

a) Whole exome sequencing in nsCLP

There is growing evidence for the involvement of rare variant having a large effect in nsCLP [181]. Previous sequencing of a few candidate genes have disclosed rare mutations in individual families [70, 182] : for example mutations in the *IRF6* gene, which is also known to contain mutations involved in a syndromic CLP, the Van der Woude syndrome (VWS).

(1) Multiplex families

Investigating multiple affected individuals from familial forms, which account for 20% of nsCLPs, will reduce heterogeneity. Several WES studies of distant relatives in multiplex families have now been conducted searching for:

- Rare coding variant in a locus previously implicated by linkage study. Investigation of the ~700 genes contained in the VWS locus (1p33-36) lead to the identification of the *GRHL3* gene as the main candidate. Large cohort sequencing and functional assay confirmed its implication in VWS and nsCLP [85].
- Rare coding variant in a list of candidate genes [183, 184].
  Unique variant in *CDH1* [183] and *MSX1* [184] were implicated in the 55 and 7 families studied respectively and few candidate genes were also suggested for further investigation [183, 184].
- Rare coding variant in the entire exome.
  WES applied in eight nsCP families, followed by resequencing in a larger cohort, have pointed out three novel candidate genes: *ACAB, PTRPS, MIBI* [185]. In 2 families, variant interpretation found in the *GRHL3* and *CREBBP* gene was limited by poor access to clinical information.
  WES studies of eight CP affected family members found a missense mutation in *ARHGAP29*, gene previously implicated in nsCLP, which impairs the protein function as shown by functional assay [186].
- Rare variant in a common gene. WES performed on 27 multiplex families found rare variant in two common genes in ≥ three families, genes which had not been implicated in clefting before: *ACSS2* and *PHYH*.

We performed WES on 79 nsCLP families and restricted our rare variants analysis to a candidate gene list. We identified mutations in 7 families, in 5 genes known to be mutated in syndromic form of CLP: *TP63, TBX1, LRP6*, *GRHL3* (2 families), *TBX22* (2 families) [187]. After clinical reassessment and confirmation of the isolated nature of the CLP, except for one of the *TBX22* mutated family, these data supported the significant overlap between syndromic and non-syndromic CLP, and the ~10% contribution to nsCLP of genes implicated in syndromic disease [25].
Very recently rare pathogenic variants were described in 4 genes in the Epithelial Cadherin-p120-catenin Complex -*CTNND1, PLEKHA7, PLEKHA5*- and a

epithelial splicing regulator -*ESRP2*-, in addition to the known CL/P associated gene *CDH1* [188].

(2) **Subgroups**

Focusing on a specific phenotype of nsCLP should reduce its clinical heterogeneity.

WES was conducted on 6 nsCP families looking for rare coding variant in common gene. One gene, *LRRC1* was taken into account for further investigation (publication IV).

We selected individuals on a distinct CLP phenotype, performing WES on eight patients with discontinuous CLP, searching for rare deleterious mutations in a candidate gene list. Mutations were found in three patients, implicating *FGFR1, ARHGAP29* and *DLG1* genes, representing almost 40 % of this cohort. Although we did not identify a unique common gene, this approach appears to be promising, as the mutation pick-up rate in known cleft genes is much higher than in familial studies (publication III).

As for others genetic tools, it must not be forgotten some of known WES limitations: technical limitations including poor coverage of some coding regions, bioinformatics variant calling issues, and variant misinterpretation [176].

b) Whole Genome sequencing in nsCLP

Targeted or whole genome sequencing extends the level of detection beyond coding regions to non-coding DNA in which specific sequences have been shown to regulate gene expression such as Cis regulatory elements, promoters and non-coding RNAs, and which are undoubtedly a key component of several craniofacial disorders. Several examples of Cis dysregulation exists in nsCLP such as a common risk variant (rs642961) altering the binding site of *TFAP2A* at the *IRF6* locus [22], or non-coding regions around SOX9 [189]. Furthermore, around 40 nsCLP associated risk regions have been identified by GWAS, the majority of which are located in non-coding regions [190].

WGS also allows detection of small deletions or larger structural changes (copy number variations), and more complex structural variants such as inversions and translocations, and study of the combined effect of common and rare variants in association studies.

Leslie and colleagues first used a “targeted sequencing” approach to study nsCLP. They chose 13 loci, encompassing ~ 6,3Mb, based on previous GWAS/genome-

wide linkage studies (9/13) and candidate gene studies (4/13). A variety of common and rare variants associated with nsCLP as well a multiple de novo mutations, with a subset occurring within predicted regulatory elements, were detected [181].

To date, only one study has been published analyzing WGS (107 affected, 6 unaffected from 32 families) in nsCLP multiplex families [191]. Analyses were mainly focused on the coding regions because of difficulties in the interpretation of the non-coding genomic regions. The authors have also pointed a major limitation for their population based analyses, as many of populations they studied were not well represented in available allele frequency databases (e.g.1000 Genomes and GnomAD).

As illustrated the use of next generation sequencing shifts disease gene identification from the identification to the interpretation phase.
Once identified, functional tests are required to determine the consequences of the specific alteration. Analysis of larger cohorts of multi-case nsCLP families and different populations are also needed to establish a possible association to nsCLP phenotype and the expanding list of genes pinpointed by WES studies. In most published studies, WES data have been made publicly available by the authors.

2. Functional approaches

Although not a functional test, it is valuable to obtain spatial and temporal information concerning the expression pattern associated with the gene under analysis as this can often provide insight into its potential role in craniofacial development. Considerable knowledge regarding gene expression pattern can now be found in several public repositories and databases (Facebase, Genepaint, Eurexpress, SysFACE …) [192].

In vitro based assays provides a convenient system to test biochemical and molecular based properties of the mutation, but only animal model provides a powerful way to dissect the role of the mutation in a functional context during development. Because the mouse is the closest animal model to human and has more closely matched anatomical structures, such as the secondary palate and teeth, this model is the “gold” standard [192]. Zebrafish has also become increasingly popular to model craniofacial diseases. In addition to the variety of knock out or knock-down technologies to disrupt gene function available in animal

model, huge advances have also been made in tools available for genome editing, particularly in model organisms. The CRISPR/Cas system has become the genome-editing tool of choice [192].

Recent implications of candidate gene detected by next generation approach followed by animal model use to further understand the consequence of it on craniofacial development concern the *GHRL3* gene (using zebrafish and mouse) [85] and the *FGFR2* gene (using zebrafish) [181].

B. nsCLP phenotyping

Accurate phenotyping is crucial to understand both the epidemiology and etiology of any congenital malformation, and the power to detect effects is strengthened when homogeneous groups are treated as a single entity.
Accurate identification of affected individuals within families is crucial when dealing with a complex genetic trait such as nsCLP, where the disease phenotype is highly variable.

1. CLP classification

Cleft lip and/or palate is phenotypically diverse, making classification difficult and challenging. Many system have been developed over the last decades [193], and are primarily divided into anatomic systems, useful for surgeons, and embryology-based systems useful for genetic counseling and research. Detailed registry systems distinguish the different affected anatomical structures: that is lip, alveolus, primary palate, secondary hard palate, and soft palate as they may each be affected in varying degrees, or not at all, the laterality and the associated morphological features.

a) CL/P versus CP

Clefts are classically divided into 2 broad categories of anatomical defect: cleft palate (CP) and cleft lip (CL) with or without cleft palate (CL/P) -CL and cleft lip and palate (CL+CP) being historically grouped into a single phenotype-. It is consistent with the distinct developmental origin of the lip/primary palate versus the secondary palate and the general observation that these 2 conditions do not segregate in the same family, although exceptions have been reported [13].

However, increasing lines of evidence suggest that there may be distinct underlying genetic causes for CL and CL+CP. These include epidemiological studies showing differences between CL and CL+CP [21] and specific recurrence risks [62]. Few studies have examined these differences in a genetic context. Previous studies have indicated that variants in *IRF6* are more strongly associated with CL than CL+CP [22], whereas *SPRY2* has demonstrated some evidence of CL+CP-specific association [194]. Similarly, analysis of the putative risk locus on human chromosome 15q13 in different nsCL/P cohorts, using detailed clinical subphenotype information, showed that *GREM1* may be specifically associated with clefts in the lip and soft palate [195]. Finally, among the several combinations of anatomical defects observed in CLP, one phenotype is referred to as "discontinuous cleft" that is the association of CL and CP with preservation of part of the primary palate [196]. Examining CL/P subtypes may help to better understand mechanisms underlying CL or CL and CP phenotype, and identify further genetic risk variant for nsCL/P.

b) Syndromic versus non-syndromic

Cleft can also be divided into 2 groups: syndromic versus non syndromic, the latter meaning that cleft does not occur in combination with other developmental, structural or cognitive anomalies. Several important CLP risk factors have been identified directly from human analyses using syndromic CLP patients, either where large Mendelian families were available or where specific additional phenotypic features allowed patients to be subdivided into more homogeneous and readily analyzable group [25]. Distinction between syndromic, which often present with various expression and incomplete penetrance, and non-syndromic forms could be difficult, and sometimes, even after a meticulous clinical examination, impossible. For example, Van der Woude syndrome (VWS, OMIM 119300) is only phenotypically distinguished from nsCLP by "minor" clinical signs such as lower pits or hypodontia, which are also variably present in VWS affected individuals -20% of the VWS patients do not have lip pits- and 10 to 15 % of VWS appears as nsCLP. Another example is X-linked cleft palate with or without ankyloglossia syndrome (CPX, OMIM 303400), associated to mutation in *TBX22* gene, in which affected females can present CP as the sole feature. Moreover ankyloglossia that could be defined as short or anteriorly attached lingual frenulum associated with limited mobility of the tongue is clinically a very subjective feature. Such a "minor" feature can therefore be missed or unreported.

We reported one nsCP family in which CPX diagnosis could only be made by involvement of the *TBX22* gene by WES sequencing (publication II).
Several syndromic genes, such *IRF6, FGFR1, TBX22, P63, MSX1* and *GHRL3,* have now been implicated in nsCLP, through variable penetrance perhaps due to the action of modifiers [25]. Further investigation of such patients could help deciphering the penetrance variability.

c) Fusion and/or differentiation defect

Various cleft types originate from different developmental time periods and have different exposition to genes and environmental factors. A classification based on human embryology of the primary and secondary palates was introduced in 2012 [197]. Cleft anomalies of the primary and secondary palate (cleft lip/alveolus, cleft lip/alveolus and palate, and cleft palate) were classified according to timing : early embryonic period 4-7weeks post conception or late embryonic period 7-12 weeks post conception; and underlying fusion and differentiation mechanisms in embryogenesis. It provided new cleft subgroups: differentiation, fusion, fusion-differentiation or fusion and differentiation defects. Almost all subphenotypes could be related to specific weeks in embryonic development [197, 198].

2. "Expanded" cleft phenotye

There is a great variability in degree of cleft malformation. It clinically ranges from an "obvious" phenotype to minimal involvements more difficult to detect including: bifid uvula, linear lip indentations with or without nostril deformity (intrauterine healed cleft), lateral upper lip fistula, and submucous palatal cleft [7].
Moreover, the range of variation may extend beyond the externally visible cleft and a wide variety of subclinical morphological variations has now been associated with overt clefts. This "expanded cleft phenotype" includes orbicularis oris muscle (OOM) defects, dental anomalies, lip dermatoglyphics, facial measurements, brain variants on MRI, velopharyngeal insufficiency (VPI) [13, 67]. The significance of associated dental anomalies remains unclear, as they can represent either microforms of orofacial clefting [199] or generalized developmental disturbances and in the end, interpretation may depend on the location, extent, and type of defect(s) present. [51].
Subclinical features are found in increased frequency in individuals with nsCLP and for some of them in their relatives [33]. They could represent the mildest

physical expression of OFC risk genes (OOM defects, VPI) and/or pleiotropic effects of risk genes (e.g.: lip dermatoglyphics) [67].

3. Reverse phenotyping

By exome sequencing of an "a priori" nsCP family, we found a causal mutation in the *TBX22* gene. Clinical assessment confirmed diagnosis of CPX and noted 2 bony overgrowths on the posterior edge of the hard palate, on each side of the palatal midline. This palatal sign cosegregates with the *TBX22* mutation. As *tbx22* is a major determinant for intramembranous bone formation in the posterior hard palate [200], such a sign raises the question of its specificity -it has never been reported in the literature before- and the underlying pathophysiology.
In addition choanal atresia was reported in 2 affected patients. This sign is consistent with the embryological expression pattern of *TBX22* and with defects seen in mouse model, expanding the phenotypic spectrum associated to *TBX22*[200].

C. Diagnostic, genetic counseling and prevention

Considering segregation studies, linkage studies, increasing data on rare variants (implicated or not in syndromic forms of CLP), together with GWAS studies, it is likely that genetic factor influencing risk to nsCLP are divided into 2 broad classes: the "causal genes" and the "genetic" risk factors ones. A few "causal genes" with "major" effects (strongly associated with risk) may simultaneously control truly Mendelian malformation syndromes and show association with apparent non-syndromic forms of orofacial clefts, while possibly many "genetic risk factors" have the typical modest individual effect on risk (as identified through GWAS). The former could, under certain circumstances, including exposure to environmental risk factors, affects the normal craniofacial development [88]. Common variant in "causal" genes could also contribute to multifactorial risk of nsCLP [201].

1. Diagnostic

CLP are being diagnosed prenatally more frequently, with an overall ultrasound detection rate of 65%, although CP is still barely detectable. In presumed isolated

CL/P or CP, prenatal invasive testing, preferably by array based methods, is recommended, and should be considered for CL [71].

Postnatal usual modalities of clinical diagnosis are used namely: family and thorough gestational history, a meticulous clinical examination, including a specific enquiry about minor clinical signs of clefts.
Molecular investigation could be proposed following the following flow chart (Figure 5).

Although classical genetics tools are still utilized, the fast development of the next generation technologies has made WES (and soon WGS?) feasible and affordable, and is likely to become a routine in diagnostic settings. For nsCLP a candidate-gene strategy applied to WES would likely be the most relevant with WES data being filtered through a fixed list of genes which have proven their definitive role in nsCLP. Being understood that such a list could evolve with biological knowledge. It would also work as a precaution from finding putatively damaging variants including those completely unrelated to the clefting phenotypes (including potential incidental findings).

As a first line assessment in multiplex families, associated CLP genes reported in Figure 5 should be tested. Because of recent identification of pathogenic variations in 5 genes in the Epithelial Cadherin-p120-Catenin Complex in 14% of nsCL/P multiplex families, testing p120-Ecadherin complex genes might also soon be taken into consideration, with *PLEKHA7* representing a strong candidate gene. [188].

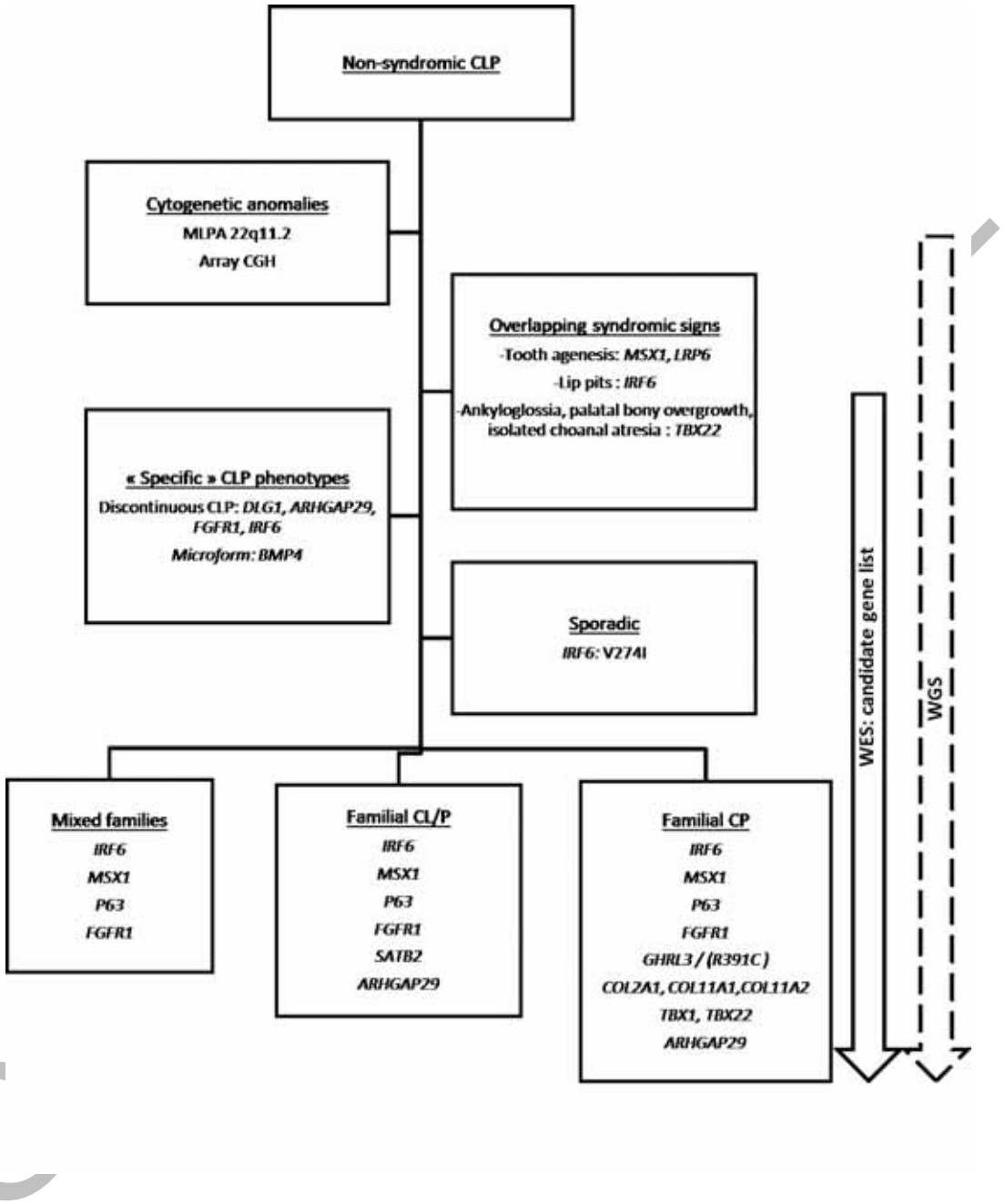


Figure 5. Non-syndromic CLP genetic investigation flow chart

2. Genetic counselling

Familial forms of nsCLP have a raised risk of CLP compared to the empirical value of three to five percent. Many familial forms follow a Mendelian pattern of transmission, and have to be counselled according to being aware of incomplete penetrance. More data has to be collected to estimate recurrence risk for each gene. If a genetic diagnosis has been made, counselling is as for the given genetic diagnosis. Because of incomplete penetrance and possibility of phenocopy (publications I and II), familial genetic testing has to be widely proposed both to affected and unaffected patient to determine their genetic status.

For most sporadic cases, empirical recurrence risk will be used according to the type of cleft. Screening of few variants known to be associated to nsCLP could be proposed to refine estimates of the risk ie V274I in *IRF6* gene [201]. Taking into consideration that *de novo* single gene variants could explain a substantial proportion of nsCL/P [188] and in order to provide a personalised rather than a population derived empiric recurrence risk, molecular testing could be discussed on a case by case basis. In case of "specific" type of nsCLP, genetic test has to be discussed in accordance to the literature (publication III).

3. Prevention

A substantial number of studies suggest that genetic and environmental risk factors may interact to modify risk of CLP, and maternal smoking, maternal alcohol consumption, and vitamin supplementation (a protective factor) have yielded statistical evidence even if results from different studies are not always consistent [61]. While maternal intake of folate and folic acid is specifically associated with decreased risk of neural tube defect, it may also provide protection for other selected congenital anomalies, including oral cleft. The current recommendation is that all women at childbearing age should take 400 ug of folic acid per day.

D. Perspectives

1. nsCLP

- Replication of the WES study on discontinuous cleft would be an asset to corroborate our data on that specific phenotype.

- "Negative" WESed family data will be further analysed with
  - An "updated" list of candidate genes,
  - Using additional combined strategies ie: entire exome + variant shared in ≥ two families
- Collaborative works on recording clinical information in the most accurate way are on-going.
- We also began to collect lip and/or palate tissue, as DNA from lymphocytes may not accurately represent the tissues affected in cleft. Occurrence of somatic mutation as the underlying mechanism of sporadic cleft has been proposed to explain the discordance between monozygotic twins for cleft [202]. Searching for discordant DNA variation between DNA extracted from tissue and blood could help to focus on interesting genes (although postnatal collected tissue may imperfectly reflect fetal tissues from early developmental stages).

2. VWS

Mutations in *IRF6* gene account for ~70% of cases with VWS, and recently grainy head-like 3 (*GRHL3*) mutations were identified in a further ~5%. VWS is the most common syndromic form of CLP, and can mimic non syndromic cleft.
We found pathogenic mutation in *LRRC1* gene in 2 distinct families, one with nsCP and one with VWS [preliminary results]. Only few data are available about that gene and interpretation of the variants needs further investigations. We could show that mutated RNA is stable, as is the mutated protein.

- We plan to screen *LRRC1* in our *IRF6* negative VWS cohort and in nsCP to search for other patients with *LRRC1* mutation. Our request on "genematcher", a freely accessible web site designed to enable connections between clinicians and researchers from around the world who share an interest in the same gene or genes, has been inconclusive until now.
- The analysis of the expression of the gene in mice, at several embryological stages on the craniofacial region would help to strengthen the involvement in palatal formation

VI. GENERAL CONCLUSION

WES is a powerful approach to identify candidate genes in selected CLP families and we could confirm the implication of syndromic CLP genes in ~10% of familial nsCLP.

Classifications based on anatomic and embryologic systems are currently used, but as additional progress is made, they become inadequate to further phenotypic research and should be refined. New classification should be more centered on gene and phenotype. Development of "phenomics" tools would allow researchers to compare and pool clinical findings in their patients and correlate them with the growing results of NGS.

While progresses have been made in identifying genetic etiologies, a biological pathway for nsCLP has not yet been established. Only few genes in different signaling pathways have been implicated in nsCLP, altogether with few upstream and downstream factors. By identifying mutations in functionally linked genes, WES technology recently allowed to highlight the epithelial cadherin-catenin complex, implicated in epithelial adhesion, as an important biological pathway for CL/P.

At the individual and population levels, understanding the factors underlying nsCLP phenotypes is important to improve prevention, treatment, prognosis of the condition and knowledge on the underlying genetic and environmental factors. At the molecular and cellular levels, "-omics studies" including genomics, epigenomics, transcriptomics, and proteomics in various cells and tissues will assist in elucidating the underlying mechanisms of these diseases.

VII. REFERENCES

Web Ressources

Online Mendelian Inheritance in Man (OMIM): http://www.omim.org
1000 Genomes: http://www.1000genomes.org/
Genome of the Netherlands (GoNL): http://www.nlgenome.nl/
ExAC: Exome Aggregation Consortium http://exac.broadinstitute.org/
Genome Aggregation Database (gnomAD): http://gnomad.broadinstitute.org/
The Human Gene Mutation Database (HMGD): http://www.hgmd.cf.ac.uk/
Expression atlas: https://www.ebi.ac.uk
Facebase: https://www.facebase.org
Genepaint: http://genepaint.org/
Eurexpress: http://www.eurexpress.org/ee/
SysFACE: https://bioinformatics.udel.edu/research/sysface/
Genematcher: https://genematcher.org

Bibliography

1. Couly, G.F., P.M. Coltey, and N.M. Le Douarin, *The developmental fate of the cephalic mesoderm in quail-chick chimeras*. Development, 1992. **114**(1): p. 1-15.
2. Minoux, M. and F.M. Rijli, *Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development*. Development, 2010. **137**(16): p. 2605-21.
3. Gaare, J.D. and J. Langman, *Fusion of nasal swellings in the mouse embryo: regression of the nasal fin.* Am J Anat, 1977. **150**(3): p. 477-99.
4. Jiang, R., J.O. Bush, and A.C. Lidral, *Development of the upper lip: morphogenetic and molecular mechanisms.* Dev Dyn, 2006. **235**(5): p. 1152-66.
5. Sun, D., S. Baur, and E.D. Hay, *Epithelial-mesenchymal transformation is the mechanism for fusion of the craniofacial primordia involved in morphogenesis of the chicken lip*. Dev Biol, 2000. **228**(2): p. 337-49.
6. Oostrom, C.A., et al., *Median cleft of the lower lip and mandible: case reports, a new embryologic hypothesis, and subdivision.* Plast Reconstr Surg, 1996. **97**(2): p. 313-20.
7. Hennekam, R.C.M., et al., *Gorlin's syndromes of the head and neck*. 5th ed. Oxford monographs on medical genetics. 2010, New York: Oxford University Press. lxi, 1452 p.

8. Ferguson, M.W., *Palate development.* Development, 1988. **103 Suppl**: p. 41-60.
9. Bush, J.O. and R. Jiang, *Palatogenesis: morphogenetic and molecular mechanisms of secondary palate development.* Development, 2012. **139**(2): p. 231-43.
10. Chiquet, M., et al., *Mesenchymal Remodeling during Palatal Shelf Elevation Revealed by Extracellular Matrix and F-Actin Expression Patterns.* Front Physiol, 2016. **7**: p. 392.
11. Vaziri Sani, F., et al., *Fate-mapping of the epithelial seam during palatal fusion rules out epithelial-mesenchymal transformation.* Dev Biol, 2005. **285**(2): p. 490-5.
12. Baek, J.A., et al., *Bmpr1a signaling plays critical roles in palatal shelf growth and palatal bone formation.* Dev Biol, 2011. **350**(2): p. 520-31.
13. Dixon, M.J., et al., *Cleft lip and palate: understanding genetic and environmental influences.* Nat Rev Genet, 2011. **12**(3): p. 167-78.
14. Tessier, P., *Anatomical classification facial, cranio-facial and latero-facial clefts.* J Maxillofac Surg, 1976. **4**(2): p. 69-92.
15. Tolarova, M.M. and J. Cervenka, *Classification and birth prevalence of orofacial clefts.* Am J Med Genet, 1998. **75**(2): p. 126-37.
16. Kondo, S., et al., *Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes.* Nat Genet, 2002. **32**(2): p. 285-9.
17. van den Boogaard, M.J., et al., *MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans.* Nat Genet, 2000. **24**(4): p. 342-3.
18. Braybrook, C., et al., *The T-box transcription factor gene TBX22 is mutated in X-linked cleft palate and ankyloglossia.* Nat Genet, 2001. **29**(2): p. 179-83.
19. Dode, C., et al., *Loss-of-function mutations in FGFR1 cause autosomal dominant Kallmann syndrome.* Nat Genet, 2003. **33**(4): p. 463-5.
20. Clements, S.E., et al., *Molecular basis of EEC (ectrodactyly, ectodermal dysplasia, clefting) syndrome: five new mutations in the DNA-binding domain of the TP63 gene and genotype-phenotype correlation.* Br J Dermatol, 2010. **162**(1): p. 201-7.
21. Harville, E.W., et al., *Cleft lip and palate versus cleft lip only: are they distinct defects?* Am J Epidemiol, 2005. **162**(5): p. 448-53.
22. Rahimov, F., et al., *Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip.* Nat Genet, 2008. **40**(11): p. 1341-7.
23. Genisca, A.E., et al., *Orofacial clefts in the National Birth Defects Prevention Study, 1997-2004.* Am J Med Genet A, 2009. **149A**(6): p. 1149-58.
24. Jones, M.C., *Etiology of facial clefts: prospective evaluation of 428 patients.* Cleft Palate J, 1988. **25**(1): p. 16-20.

25. Stanier, P. and G.E. Moore, *Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts.* Hum Mol Genet, 2004. **13 Spec No 1**: p. R73-81.
26. Bush, P.G. and A.J. Williams, *Incidence of the Robin Anomalad (Pierre Robin syndrome).* Br J Plast Surg, 1983. **36**(4): p. 434-7.
27. Printzlau, A. and M. Andersen, *Pierre Robin sequence in Denmark: a retrospective population-based epidemiological study.* Cleft Palate Craniofac J, 2004. **41**(1): p. 47-52.
28. Abadie, V., et al., *Brainstem dysfunction: a possible neuroembryological pathogenesis of isolated Pierre Robin sequence.* Eur J Pediatr, 2002. **161**(5): p. 275-80.
29. Latham, R.A., *The pathogenesis of cleft palate associated with the Pierre Robin syndrome. An analysis of a seventeen-week human foetus.* Br J Plast Surg, 1966. **19**(3): p. 205-14.
30. Marques, I.L., M.A. Barbieri, and H. Bettiol, *Etiopathogenesis of isolated Robin sequence.* Cleft Palate Craniofac J, 1998. **35**(6): p. 517-25.
31. Holder-Espinasse, M., et al., *Pierre Robin sequence: a series of 117 consecutive cases.* J Pediatr, 2001. **139**(4): p. 588-90.
32. Shprintzen, R.J., *Pierre Robin, micrognathia, and airway obstruction: the dependency of treatment on accurate diagnosis.* Int Anesthesiol Clin, 1988. **26**(1): p. 64-71.
33. Jugessur, A., P.G. Farlie, and N. Kilpatrick, *The genetics of isolated orofacial clefts: from genotypes to subphenotypes.* Oral Dis, 2009. **15**(7): p. 437-53.
34. Suzuki, S., et al., *Mutations in BMP4 are associated with subepithelial, microform, and overt cleft lip.* Am J Hum Genet, 2009. **84**(3): p. 406-11.
35. Neiswanger, K., et al., *Whorl patterns on the lower lip are associated with nonsyndromic cleft lip with or without cleft palate.* Am J Med Genet A, 2009. **149A**(12): p. 2673-9.
36. Letra, A., et al., *Defining subphenotypes for oral clefts based on dental development.* J Dent Res, 2007. **86**(10): p. 986-91.
37. Lourenco Ribeiro, L., et al., *Dental anomalies of the permanent lateral incisors and prevalence of hypodontia outside the cleft area in complete unilateral cleft lip and palate.* Cleft Palate Craniofac J, 2003. **40**(2): p. 172-5.
38. Eerens, K., et al., *Hypodontia and tooth formation in groups of children with cleft, siblings without cleft, and nonrelated controls.* Cleft Palate Craniofac J, 2001. **38**(4): p. 374-8.
39. Peters, H., A. Neubuser, and R. Balling, *Pax genes and organogenesis: Pax9 meets tooth development.* Eur J Oral Sci, 1998. **106 Suppl 1**: p. 38-43.

40. Vieira, A.R., et al., *Candidate gene/loci studies in cleft lip/palate and dental anomalies finds novel susceptibility genes for clefts.* Genet Med, 2008. **10**(9): p. 668-74.
41. Vieira, A.R., et al., *MSX1, PAX9, and TGFA contribute to tooth agenesis in humans.* J Dent Res, 2004. **83**(9): p. 723-7.
42. De Mey, A., et al., *Anatomy of the orbicularis oris muscle in cleft lip.* Br J Plast Surg, 1989. **42**(6): p. 710-4.
43. Neiswanger, K., et al., *Orbicularis oris muscle defects as an expanded phenotypic feature in nonsyndromic cleft lip with or without cleft palate.* Am J Med Genet A, 2007. **143A**(11): p. 1143-9.
44. Martin, R.A., et al., *Ultrasonographic detection of orbicularis oris defects in first degree relatives of isolated cleft lip patients.* Am J Med Genet, 2000. **90**(2): p. 155-61.
45. Schmidt, K.L., et al., *Nasolabial fold discontinuity during speech as a possible extended cleft phenotype.* Cleft Palate Craniofac J, 2013. **50**(2): p. 201-6.
46. Roosenboom, J., et al., *Review: Facial endophenotypes in non-syndromic orofacial clefting.* B-ENT, 2015. **11**(3): p. 173-82.
47. Trasler, D.G., *Pathogenesis of cleft lip and its relation to embryonic face shape in A-J and C57BL mice.* Teratology, 1968. **1**(1): p. 33-49.
48. Fraser, F.C. and H. Pashayan, *Relation of face shape to susceptibility to congenital cleft lip. A preliminary report.* J Med Genet, 1970. **7**(2): p. 112-7.
49. Claes, P., M. Walters, and J. Clement, *Improved facial outcome assessment using a 3D anthropometric mask.* Int J Oral Maxillofac Surg, 2012. **41**(3): p. 324-30.
50. Weinberg, S.M., et al., *Face shape of unaffected parents with cleft affected offspring: combining three-dimensional surface imaging and geometric morphometrics.* Orthod Craniofac Res, 2009. **12**(4): p. 271-81.
51. Weinberg, S.M., et al., *The Pittsburgh Oral-Facial Cleft study: expanding the cleft phenotype. Background and justification.* Cleft Palate Craniofac J, 2006. **43**(1): p. 7-20.
52. Roosenboom, J., et al., *Testing the face shape hypothesis in twins discordant for nonsyndromic orofacial clefting.* Am J Med Genet A, 2017. **173**(11): p. 2886-2892.
53. Boehringer, S., et al., *Genetic determination of human facial morphology: links between cleft-lips and normal variation.* Eur J Hum Genet, 2011. **19**(11): p. 1192-7.
54. Conrad, A.L., et al., *Incidence of neurological soft signs in children with isolated cleft of the lip or palate.* Percept Mot Skills, 2008. **106**(1): p. 197-206.
55. Nopoulos, P., et al., *Structural brain abnormalities in adult males with clefts of the lip and/or palate.* Genet Med, 2002. **4**(1): p. 1-9.

56. Nopoulos, P., et al., *Increased incidence of a midline brain anomaly in patients with nonsyndromic clefts of the lip and/or palate.* J Neuroimaging, 2001. **11**(4): p. 418-24.
57. Conrad, A.L., et al., *Cerebellum structure differences and relationship to speech in boys and girls with nonsyndromic cleft of the lip and/or palate.* Cleft Palate Craniofac J, 2010. **47**(5): p. 469-75.
58. Nopoulos, P., et al., *Abnormal brain structure in adults with Van der Woude syndrome.* Clin Genet, 2007. **71**(6): p. 511-7.
59. May, M.A., et al., *Evidence of olfactory deficits as part of the phenotypic spectrum of nonsyndromic orofacial clefting.* J Craniofac Surg, 2015. **26**(1): p. 84-6.
60. Roosenboom, J., et al., *Facial Characteristics and Olfactory Dysfunction: Two Endophenotypes Related to Nonsyndromic Cleft Lip and/or Palate.* Biomed Res Int, 2015. **2015**: p. 863429.
61. Mossey, P.A., et al., *Cleft lip and palate.* Lancet, 2009. **374**(9703): p. 1773-85.
62. Grosen, D., et al., *A cohort study of recurrence patterns among more than 54,000 relatives of oral cleft cases in Denmark: support for the multifactorial threshold model of inheritance.* J Med Genet, 2010. **47**(3): p. 162-8.
63. Mitchell, L.E. and N. Risch, *Correlates of genetic risk for non-syndromic cleft lip with or without cleft palate.* Clin Genet, 1993. **43**(5): p. 255-60.
64. Fogh-Andersen, P., *Inheritance of harelip and cleft palate; contribution to the elucidation of the etiology of the congenital clefts of the face*. 1942: Copenhagen,. p. 266 p.
65. Grosen, D., et al., *Risk of oral clefts in twins.* Epidemiology, 2011. **22**(3): p. 313-9.
66. Gajdos, V., et al., *Genetics of nonsyndromic cleft lip with or without cleft palate: is there a Mendelian sub-entity?* Ann Genet, 2004. **47**(1): p. 29-39.
67. Marazita, M.L., *The evolution of human genetic studies of cleft lip and cleft palate.* Annu Rev Genomics Hum Genet, 2012. **13**: p. 263-83.
68. Christensen, K. and L.E. Mitchell, *Familial recurrence-pattern analysis of nonsyndromic isolated cleft palate--a Danish Registry study.* Am J Hum Genet, 1996. **58**(1): p. 182-90.
69. Schliekelman, P. and M. Slatkin, *Multiplex relative risk and estimation of the number of loci underlying an inherited disease.* Am J Hum Genet, 2002. **71**(6): p. 1369-85.
70. Riley, B.M., et al., *Impaired FGF signaling contributes to cleft lip and palate.* Proc Natl Acad Sci U S A, 2007. **104**(11): p. 4512-7.
71. Maarse, W., et al., *A systematic review of associated structural and chromosomal defects in oral clefts: when is prenatal genetic analysis indicated?* J Med Genet, 2012. **49**(8): p. 490-8.

72. FitzPatrick, D.R., et al., *Identification of SATB2 as the cleft palate gene on 2q32-q33.* Hum Mol Genet, 2003. **12**(19): p. 2491-501.
73. Alkuraya, F.S., et al., *SUMO1 haploinsufficiency leads to cleft lip and palate.* Science, 2006. **313**(5794): p. 1751.
74. Kim, H.G., et al., *Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for FGFR1.* J Med Genet, 2005. **42**(8): p. 666-72.
75. Jia, Z.L., et al., *Interactions between small ubiquitin-like modifier 1 and nonsyndromic orofacial clefts.* DNA Cell Biol, 2011. **30**(4): p. 235-40.
76. Shi, M., et al., *Identification of microdeletions in candidate genes for cleft lip and/or palate.* Birth Defects Res A Clin Mol Teratol, 2009. **85**(1): p. 42-51.
77. Ghassibe-Sabbagh, M., et al., *FAF1, a gene that is disrupted in cleft palate and has conserved function in zebrafish.* Am J Hum Genet, 2011. **88**(2): p. 150-61.
78. Sahoo, T., et al., *Microdeletion 20p12.3 involving BMP2 contributes to syndromic forms of cleft palate.* Am J Med Genet A, 2011. **155A**(7): p. 1646-53.
79. Osoegawa, K., et al., *Identification of novel candidate genes associated with cleft lip and palate using array comparative genomic hybridisation.* J Med Genet, 2008. **45**(2): p. 81-6.
80. Conte, F., et al., *Systematic analysis of copy number variants of a large cohort of orofacial cleft patients identifies candidate genes for orofacial clefts.* Hum Genet, 2016. **135**(1): p. 41-59.
81. Riley, B.M. and J.C. Murray, *Sequence evaluation of FGF and FGFR gene conserved non-coding elements in non-syndromic cleft lip and palate cases.* Am J Med Genet A, 2007. **143A**(24): p. 3228-34.
82. Ng, S.B., et al., *Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome.* Nat Genet, 2010. **42**(9): p. 790-3.
83. Ng, S.B., et al., *Exome sequencing identifies the cause of a mendelian disorder.* Nat Genet, 2010. **42**(1): p. 30-5.
84. Mitchell, K., et al., *Exome sequence identifies RIPK4 as the Bartsocas-Papas syndrome locus.* Am J Hum Genet, 2012. **90**(1): p. 69-75.
85. Peyrard-Janvid, M., et al., *Dominant mutations in GRHL3 cause Van der Woude Syndrome and disrupt oral periderm development.* Am J Hum Genet, 2014. **94**(1): p. 23-32.
86. Satokata, I. and R. Maas, *Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development.* Nat Genet, 1994. **6**(4): p. 348-56.
87. Koillinen, H., et al., *A genome-wide scan of non-syndromic cleft palate only (CPO) in Finnish multiplex families.* J Med Genet, 2005. **42**(2): p. 177-84.

88. Beaty, T.H., M.L. Marazita, and E.J. Leslie, *Genetic factors influencing risk to orofacial clefts: today's challenges and tomorrow's opportunities.* F1000Res, 2016. **5**: p. 2800.
89. Stanier, P. and E. Pauws, *Development of the lip and palate: FGF signalling.* Front Oral Biol, 2012. **16**: p. 71-80.
90. Gritli-Linde, A., *The mouse as a developmental model for cleft lip and palate research.* Front Oral Biol, 2012. **16**: p. 32-51.
91. Smith, T.M., et al., *Molecular signaling along the anterior-posterior axis of early palate development.* Front Physiol, 2012. **3**: p. 488.
92. Weng, M., et al., *A review of FGF signaling in palate development.* Biomed Pharmacother, 2018. **103**: p. 240-247.
93. He, F. and Y. Chen, *Wnt signaling in lip and palate development.* Front Oral Biol, 2012. **16**: p. 81-90.
94. Lan, Y., et al., *Expression of Wnt9b and activation of canonical Wnt signaling during midfacial morphogenesis in mice.* Dev Dyn, 2006. **235**(5): p. 1448-54.
95. Jin, Y.R., et al., *Wnt9b-dependent FGF signaling is crucial for outgrowth of the nasal and maxillary processes during upper jaw and lip development.* Development, 2012. **139**(10): p. 1821-30.
96. Song, L., et al., *Lrp6-mediated canonical Wnt signaling is required for lip formation and fusion.* Development, 2009. **136**(18): p. 3161-71.
97. He, F., et al., *Epithelial Wnt/beta-catenin signaling regulates palatal shelf fusion through regulation of Tgfbeta3 expression.* Dev Biol, 2011. **350**(2): p. 511-9.
98. Brugmann, S.A., et al., *Wnt signaling mediates regional specification in the vertebrate face.* Development, 2007. **134**(18): p. 3283-95.
99. Lin, C., et al., *The inductive role of Wnt-beta-Catenin signaling in the formation of oral apparatus.* Dev Biol, 2011. **356**(1): p. 40-50.
100. Schwabe, G.C., et al., *Ror2 knockout mouse as a model for the developmental pathology of autosomal recessive Robinow syndrome.* Dev Dyn, 2004. **229**(2): p. 400-10.
101. He, F., et al., *Wnt5a regulates directional cell migration and cell proliferation via Ror2-mediated noncanonical pathway in mammalian palate development.* Development, 2008. **135**(23): p. 3871-9.
102. Lee, J.M., et al., *Wnt11/Fgfr1b cross-talk modulates the fate of cells in palate development.* Dev Biol, 2008. **314**(2): p. 341-50.
103. Marazita, M.L., et al., *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): p. 161-73.
104. Menezes, R., et al., *Studies with Wnt genes and nonsyndromic cleft lip and palate.* Birth Defects Res A Clin Mol Teratol, 2010. **88**(11): p. 995-1000.

105. Chiquet, B.T., et al., *Variation in WNT genes is associated with non-syndromic cleft lip with or without cleft palate.* Hum Mol Genet, 2008. **17**(14): p. 2212-8.
106. Niemann, S., et al., *Homozygous WNT3 mutation causes tetra-amelia in a large consanguineous family.* Am J Hum Genet, 2004. **74**(3): p. 558-63.
107. Rice, R., et al., *Disruption of Fgf10/Fgfr2b-coordinated epithelial-mesenchymal interactions causes cleft palate.* J Clin Invest, 2004. **113**(12): p. 1692-700.
108. Lan, Y. and R. Jiang, *Sonic hedgehog signaling regulates reciprocal epithelial-mesenchymal interactions controlling palatal outgrowth.* Development, 2009. **136**(8): p. 1387-96.
109. Cobourne, M.T. and J.B. Green, *Hedgehog signalling in development of the secondary palate.* Front Oral Biol, 2012. **16**: p. 52-9.
110. Solomon, B.D., et al., *Analysis of genotype-phenotype correlations in human holoprosencephaly.* Am J Med Genet C Semin Med Genet, 2010. **154C**(1): p. 133-41.
111. Mansilla, M.A., et al., *Contributions of PTCH gene variants to isolated cleft lip and palate.* Cleft Palate Craniofac J, 2006. **43**(1): p. 21-9.
112. Shi, Y. and J. Massague, *Mechanisms of TGF-beta signaling from cell membrane to the nucleus.* Cell, 2003. **113**(6): p. 685-700.
113. Parada, C. and Y. Chai, *Roles of BMP signaling pathway in lip and palate development.* Front Oral Biol, 2012. **16**: p. 60-70.
114. Hu, L., et al., *TGFbeta3 regulates periderm removal through DeltaNp63 in the developing palate.* J Cell Physiol, 2015. **230**(6): p. 1212-25.
115. Li, C., Y. Lan, and R. Jiang, *Molecular and Cellular Mechanisms of Palate Development.* J Dent Res, 2017. **96**(11): p. 1184-1191.
116. Richardson, R.J., et al., *Irf6 is a key determinant of the keratinocyte proliferation-differentiation switch.* Nat Genet, 2006. **38**(11): p. 1329-34.
117. Ingraham, C.R., et al., *Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6).* Nat Genet, 2006. **38**(11): p. 1335-40.
118. Iwata, J., et al., *Smad4-Irf6 genetic interaction and TGFbeta-mediated IRF6 signaling cascade are crucial for palatal fusion in mice.* Development, 2013. **140**(6): p. 1220-30.
119. !!! INVALID CITATION !!! {}.
120. Lane, J., et al., *Tak1, Smad4 and Trim33 redundantly mediate TGF-beta3 signaling during palate development.* Dev Biol, 2015. **398**(2): p. 231-41.
121. Thomason, H.A., et al., *Cooperation between the transcription factors p63 and IRF6 is essential to prevent cleft palate in mice.* J Clin Invest, 2010. **120**(5): p. 1561-9.
122. Moretti, F., et al., *A regulatory feedback loop involving p63 and IRF6 links the pathogenesis of 2 genetically different human ectodermal dysplasias.* J Clin Invest, 2010. **120**(5): p. 1570-7.

123. Rinne, T., H.G. Brunner, and H. van Bokhoven, *p63-associated disorders.* Cell Cycle, 2007. **6**(3): p. 262-8.
124. Liu, W., et al., *Distinct functions for Bmp signaling in lip and palate fusion in mice.* Development, 2005. **132**(6): p. 1453-61.
125. Lin, J.Y., et al., *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): p. 601-5.
126. Jezewski, P.A., et al., *Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate.* J Med Genet, 2003. **40**(6): p. 399-407.
127. Little, J., A. Cardy, and R.G. Munger, *Tobacco smoking and oral clefts: a meta-analysis.* Bull World Health Organ, 2004. **82**(3): p. 213-8.
128. Lyons, D.C., *Biochemical factors in the possible developments of malformations of the oral tissues, especially cleft palate and harelip.* Oral Surg Oral Med Oral Pathol, 1950. **3**(1): p. 12-21, illust.
129. Holmberg, P.C., et al., *Oral clefts and organic solvent exposure during pregnancy.* Int Arch Occup Environ Health, 1982. **50**(4): p. 371-6.
130. Mossey, P.A., J.A. Davies, and J. Little, *Prevention of orofacial clefts: does pregnancy planning have a role?* Cleft Palate Craniofac J, 2007. **44**(3): p. 244-50.
131. Honein, M.A., et al., *Maternal smoking and environmental tobacco smoke exposure and the risk of orofacial clefts.* Epidemiology, 2007. **18**(2): p. 226-33.
132. Sabbagh, H.J., et al., *Passive smoking in the etiology of non-syndromic orofacial clefts: a systematic review and meta-analysis.* PLoS One, 2015. **10**(3): p. e0116963.
133. Ozturk, F., et al., *Nicotine Exposure During Pregnancy Results in Persistent Midline Epithelial Seam With Improper Palatal Fusion.* Nicotine Tob Res, 2016. **18**(5): p. 604-12.
134. Baroni, T., et al., *Human cleft lip and palate fibroblasts and normal nicotine-treated fibroblasts show altered in vitro expressions of genes related to molecular signaling pathways and extracellular matrix metabolism.* J Cell Physiol, 2010. **222**(3): p. 748-56.
135. Romitti, P.A., et al., *Maternal periconceptional alcohol consumption and risk of orofacial clefts.* Am J Epidemiol, 2007. **166**(7): p. 775-85.
136. DeRoo, L.A., et al., *First-trimester maternal alcohol consumption and the risk of infant oral clefts in Norway: a population-based case-control study.* Am J Epidemiol, 2008. **168**(6): p. 638-46.
137. Boyles, A.L., et al., *Maternal alcohol consumption, alcohol metabolism genes, and the risk of oral clefts: a population-based case-control study in Norway, 1996-2001.* Am J Epidemiol, 2010. **172**(8): p. 924-31.
138. DeRoo, L.A., et al., *Maternal alcohol binge-drinking in the first trimester and the risk of orofacial clefts in offspring: a large population-based pooling study.* Eur J Epidemiol, 2016. **31**(10): p. 1021-1034.

139. Johnson, C.Y. and J. Little, *Folate intake, markers of folate status and oral clefts: is the evidence converging?* Int J Epidemiol, 2008. **37**(5): p. 1041-58.
140. Reynolds, P.R., G.B. Schaalje, and R.E. Seegmiller, *Combination therapy with folic acid and methionine in the prevention of retinoic acid-induced cleft palate in mice.* Birth Defects Res A Clin Mol Teratol, 2003. **67**(3): p. 168-73.
141. Wehby, G.L. and J.C. Murray, *Folic acid and orofacial clefts: a review of the evidence.* Oral Dis, 2010. **16**(1): p. 11-9.
142. Krapels, I.P., et al., *Maternal dietary B vitamin intake, other than folate, and the association with orofacial cleft in the offspring.* Eur J Nutr, 2004. **43**(1): p. 7-14.
143. Krapels, I.P., et al., *Myo-inositol, glucose and zinc status as risk factors for non-syndromic cleft lip with or without cleft palate in offspring: a case-control study*. BJOG, 2004. **111**(7): p. 661-8.
144. Munger, R.G., et al., *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): p. 151-5.
145. Krapels, I.P., et al., *Maternal nutritional status and the risk for orofacial cleft offspring in humans.* J Nutr, 2004. **134**(11): p. 3106-13.
146. Johansen, A.M., et al., *Maternal consumption of coffee and caffeine-containing beverages and oral clefts: a population-based case-control study in Norway*. Am J Epidemiol, 2009. **169**(10): p. 1216-22.
147. Collier, S.A., et al., *Maternal caffeine intake during pregnancy and orofacial clefts.* Birth Defects Res A Clin Mol Teratol, 2009. **85**(10): p. 842-9.
148. Shaw, G.M., et al., *Orofacial clefts and maternal anticonvulsant use.* Reprod Toxicol, 1995. **9**(1): p. 97-8.
149. Hernandez-Diaz, S., et al., *Folic acid antagonists during pregnancy and the risk of birth defects.* N Engl J Med, 2000. **343**(22): p. 1608-14.
150. Jentink, J., et al., *Valproic acid monotherapy in pregnancy and major congenital malformations.* N Engl J Med, 2010. **362**(23): p. 2185-93.
151. Park-Wyllie, L., et al., *Birth defects after maternal exposure to corticosteroids: prospective cohort study and meta-analysis of epidemiological studies*. Teratology, 2000. **62**(6): p. 385-92.
152. Hviid, A. and D. Molgaard-Nielsen, *Corticosteroid use during pregnancy and risk of orofacial clefts*. CMAJ, 2011. **183**(7): p. 796-804.
153. Skuladottir, H., et al., *Corticosteroid use and risk of orofacial clefts.* Birth Defects Res A Clin Mol Teratol, 2014. **100**(6): p. 499-506.
154. Correa, A., et al., *Diabetes mellitus and birth defects*. Am J Obstet Gynecol, 2008. **199**(3): p. 237 e1-9.

155. Stott-Miller, M., et al., *Increased risk of orofacial clefts associated with maternal obesity: case-control study and Monte Carlo-based bias analysis.* Paediatr Perinat Epidemiol, 2010. **24**(5): p. 502-12.
156. Muraskas, J.K., et al., *Amniotic band syndrome with significant orofacial clefts and disruptions and distortions of craniofacial structures.* J Pediatr Surg, 2003. **38**(4): p. 635-8.
157. Acs, N., et al., *Maternal influenza during pregnancy and risk of congenital abnormalities in offspring.* Birth Defects Res A Clin Mol Teratol, 2005. **73**(12): p. 989-96.
158. Shahrukh Hashmi, S., et al., *Maternal fever during early pregnancy and the risk of oral clefts.* Birth Defects Res A Clin Mol Teratol, 2010. **88**(3): p. 186-94.
159. Waller, D.K., et al., *Maternal report of fever from cold or flu during early pregnancy and the risk for noncardiac birth defects, National Birth Defects Prevention Study, 1997-2011.* Birth Defects Res, 2018. **110**(4): p. 342-351.
160. Shaw, G.M., et al., *Maternal pesticide exposure from multiple sources and selected congenital anomalies.* Epidemiology, 1999. **10**(1): p. 60-6.
161. Romitti, P.A., et al., *Meta-analysis: pesticides and orofacial clefts.* Cleft Palate Craniofac J, 2007. **44**(4): p. 358-65.
162. van Rooij, I.A., et al., *Smoking, genetic polymorphisms in biotransformation enzymes, and nonsyndromic oral clefting: a gene-environment interaction.* Epidemiology, 2001. **12**(5): p. 502-7.
163. Shi, M., et al., *Orofacial cleft risk is increased with maternal smoking and specific detoxification-gene variants.* Am J Hum Genet, 2007. **80**(1): p. 76-90.
164. Shaw, G.M., et al., *Endothelial nitric oxide synthase (NOS3) genetic variants, maternal smoking, vitamin use, and risk of human orofacial clefts.* Am J Epidemiol, 2005. **162**(12): p. 1207-14.
165. Lammer, E.J., et al., *Maternal smoking and the risk of orofacial clefts: Susceptibility with NAT1 and NAT2 polymorphisms.* Epidemiology, 2004. **15**(2): p. 150-6.
166. Maestri, N.E., et al., *Application of transmission disequilibrium tests to nonsyndromic oral clefts: including candidate genes and environmental exposures in the models.* Am J Med Genet, 1997. **73**(3): p. 337-44.
167. Jianyan, L., et al., *Analysis of interactions between genetic variants of BMP4 and environmental factors with nonsyndromic cleft lip with or without cleft palate susceptibility.* Int J Oral Maxillofac Surg, 2010. **39**(1): p. 50-6.
168. Wang, H., et al., *The FGF and FGFR Gene Family and Risk of Cleft Lip With or Without Cleft Palate.* Cleft Palate Craniofac J, 2013. **50**(1): p. 96-103.

169. Wu, T., et al., *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): p. 401-10.
170. Romitti, P.A., et al., *Candidate genes for nonsyndromic cleft lip and palate and maternal cigarette smoking and alcohol consumption: evaluation of genotype-environment interactions from a population-based case-control study of orofacial clefts.* Teratology, 1999. **59**(1): p. 39-50.
171. Jia, Z.L., et al., *Association among polymorphisms at MYH9, environmental factors, and nonsyndromic orofacial clefts in western China.* DNA Cell Biol, 2010. **29**(1): p. 25-32.
172. Wu, T., et al., *Evidence of gene-environment interaction for the RUNX2 gene and environmental tobacco smoke in controlling the risk of cleft lip with/without cleft palate.* Birth Defects Res A Clin Mol Teratol, 2012. **94**(2): p. 76-83.
173. Hwang, S.J., et al., *Association study of transforming growth factor alpha (TGF alpha) TaqI polymorphism and oral clefts: indication of gene-environment interaction in a population-based sample of infants with birth defects.* Am J Epidemiol, 1995. **141**(7): p. 629-36.
174. van Rooij, I.A., et al., *Does the interaction between maternal folate intake and the methylenetetrahydrofolate reductase polymorphisms affect the risk of cleft lip with or without cleft palate?* Am J Epidemiol, 2003. **157**(7): p. 583-91.
175. Wu, J., et al., *Significant evidence of association between polymorphisms in ZNF533, environmental factors, and nonsyndromic orofacial clefts in the Western Han Chinese population.* DNA Cell Biol, 2011. **30**(1): p. 47-54.
176. Gilissen, C., et al., *Disease gene identification strategies for exome sequencing.* Eur J Hum Genet, 2012. **20**(5): p. 490-7.
177. Leslie, E.J. and J.C. Murray, *Evaluating rare coding variants as contributing causes to non-syndromic cleft lip and palate.* Clin Genet, 2013. **84**(5): p. 496-500.
178. Kousa, Y.A., et al., *Shared molecular networks in orofacial and neural tube development.* Birth Defects Res, 2017. **109**(2): p. 169-179.
179. Lough, K.J., et al., *Closing the Gap: Mouse Models to Study Adhesion in Secondary Palatogenesis.* J Dent Res, 2017. **96**(11): p. 1210-1220.
180. Leslie, E.J., et al., *Expression and mutation analyses implicate ARHGAP29 as the etiologic gene for the cleft lip with or without cleft palate locus identified by genome-wide association on chromosome 1p22.* Birth Defects Res A Clin Mol Teratol, 2012. **94**(11): p. 934-42.
181. Leslie, E.J., et al., *Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci.* Am J Hum Genet, 2015. **96**(3): p. 397-411.

182. Vieira, A.R., et al., *Medical sequencing of candidate genes for nonsyndromic cleft lip and palate.* PLoS Genet, 2005. **1**(6): p. e64.
183. Bureau, A., et al., *Whole exome sequencing of distant relatives in multiplex families implicates rare variants in candidate genes for oral clefts.* Genetics, 2014. **197**(3): p. 1039-44.
184. Pengelly, R.J., et al., *Deleterious coding variants in multi-case families with non-syndromic cleft lip and/or palate phenotypes.* Sci Rep, 2016. **6**: p. 30457.
185. Hoebel, A.K., et al., *Candidate Genes for Nonsyndromic Cleft Palate Detected by Exome Sequencing.* J Dent Res, 2017. **96**(11): p. 1314-1321.
186. Liu, H., et al., *Exome sequencing provides additional evidence for the involvement of ARHGAP29 in Mendelian orofacial clefting and extends the phenotypic spectrum to isolated cleft palate.* Birth Defects Res, 2017. **109**(1): p. 27-37.
187. Basha, M., et al., *Whole exome sequencing identifies mutations in 10% of patients with familial non-syndromic cleft lip and/or palate in genes mutated in well-known syndromes.* J Med Genet, 2018.
188. Cox, L.L., et al., *Mutations in the Epithelial Cadherin-p120-Catenin Complex Cause Mendelian Non-Syndromic Cleft Lip with or without Cleft Palate.* Am J Hum Genet, 2018. **102**(6): p. 1143-1157.
189. Benko, S., et al., *Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence.* Nat Genet, 2009. **41**(3): p. 359-64.
190. Thieme, F. and K.U. Ludwig, *The Role of Noncoding Genetic Variation in Isolated Orofacial Clefts.* J Dent Res, 2017. **96**(11): p. 1238-1247.
191. Holzinger, E.R., et al., *Analysis of sequence data to identify potential risk variants for oral clefts in multiplex families.* Mol Genet Genomic Med, 2017. **5**(5): p. 570-579.
192. Van Otterloo, E., T. Williams, and K.B. Artinger, *The old and new face of craniofacial research: How animal models inform human craniofacial genetic and clinical data.* Dev Biol, 2016. **415**(2): p. 171-187.
193. Allori, A.C., et al., *Classification of Cleft Lip/Palate: Then and Now.* Cleft Palate Craniofac J, 2017. **54**(2): p. 175-188.
194. Ludwig, K.U., et al., *Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci.* Nat Genet, 2012. **44**(9): p. 968-71.
195. Ludwig, K.U., et al., *Meta-analysis Reveals Genome-Wide Significance at 15q13 for Nonsyndromic Clefting of Both the Lip and the Palate, and Functional Analyses Implicate GREM1 As a Plausible Causative Gene.* PLoS Genet, 2016. **12**(3): p. e1005914.
196. Alvarez, C.W., M.L. Guion-Almeida, and A. Richieri-Costa, *Clinical and genetic study on 356 Brazilian patients with a distinct phenotype of cleft lip*

and palate without alveolar ridge involvement. J Craniomaxillofac Surg, 2014. **42**(8): p. 1952-7.

197. Luijsterburg, A.J., A.M. Rozendaal, and C. Vermeij-Keers, *Classifying common oral clefts: a new approach after descriptive registration.* Cleft Palate Craniofac J, 2014. **51**(4): p. 381-91.
198. Vermeij-Keers, C., et al., *Subphenotyping and Classification of Cleft Lip and Alveolus in Adult Unoperated Patients: A New Embryological Approach.* Cleft Palate Craniofac J, 2018: p. 1055665618767106.
199. Phan, M., et al., *Tooth agenesis and orofacial clefting: genetic brothers in arms?* Hum Genet, 2016. **135**(12): p. 1299-1327.
200. Pauws, E., et al., *Tbx22null mice have a submucous cleft palate due to reduced palatal bone formation and also display ankyloglossia and choanal atresia phenotypes.* Hum Mol Genet, 2009. **18**(21): p. 4171-9.
201. Zucchero, T.M., et al., *Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate.* N Engl J Med, 2004. **351**(8): p. 769-80.
202. Kimani, J.W., et al., *Search for genomic alterations in monozygotic twins discordant for cleft lip and/or palate.* Twin Res Hum Genet, 2009. **12**(5): p. 462-8.

DEMEER Bénédicte,

Next generation sequencing increases precise diagnoses in non-syndromic cleft lip and/or palate

Caractérisation clinico-moléculaire des fentes labiales et/ou palatines par séquençage haut-débit

Doctoral Thesis in Biomedical & Pharmaceutical Sciences
Thèse de Doctorat en Sciences Biologie & Santé
Promoters: Professors Bernard DEVAUCHELLE & Miikka VIKKULA

Cleft of the lip and/or palate (CLP) are among the most common birth defects. Affected children need multidisciplinary treatment and care from birth until adulthood. They most commonly occur as an isolated defect and have a complex and heterogeneous etiology with genetic predisposition as well as environmental factors acting in concert. Whole Exome Sequencing (WES) has now become a useful approach for proving clinical molecular diagnosis. We conducted WES in 79 non-syndromic (ns) CLP multiplex families to identify rare causing variants. Among families studied, seven were found to be mutated (~10%) in syndromic genes: *TP63, TBX1, LRP6, GRHL3* and *TBX22* providing further evidence of contribution of syndromic CLP genes in ns forms. Unraveling *TBX22* mutations in 2 families allowed us to describe a new clinical oral sign: bony overgrowths on the posterior edge of the hard palate, on each side of the palatal midline, as well as to expand the *TBX22*-related phenotypic spectrum to choanal atresia and Pierre-Robin sequence. We also described mutations in a new gene, *LRRC1* in 2 distinct families: one with Van der Woude syndrome and one with ns cleft palate. We then focused on a distinguishable CLP phenotype and performed WES on eight sporadic or familial patients with "discontinuous cleft" which combined a cleft of the primary and the secondary palate with preservation of part of the intermaxillary palate. Three causative mutations were found in genes previously implicated in human CLP phenotype and/or syndrome. Our work illustrates how well-defined phenotype combined with the power of high throughput technologies significantly contribute to the understanding of genetic factor underlying CLP, and how it could help more accurate clinical, epidemiological, and fundamental research, ultimately resulting in better diagnosis and care of CLP patients.

Les fentes labiales et/ou palatines (FLP) représentent les plus fréquentes des malformations congénitales faciales. Elles requièrent une prise en charge multidisciplinaire au long terme. Elles se produisent majoritairement de façon isolée. Leur cause est complexe et hétérogène, associant des interactions entre prédisposition génétique et facteurs environnementaux. Parmi les nouvelles technologies d'étude du génome, le séquençage haut débit d'exome (WES) est devenu un outil diagnostique moléculaire majeur. Nous avons étudié par WES 79 familles multiplexes porteuses de FLP non-syndromique pour identifier des variants rares causaux. Chez sept d'entre elles, soit chez environ 10%, une mutation dans un gène impliqué dans les FLPs syndromiques ont été mises en évidence : *TP63, TBX1, LRP6, GRHL3* et *TBX22* renforçant la notion de contribution de certains gènes communs aux formes syndromiques et non syndromiques. La réévaluation clinique des 2 familles porteuses de mutation dans *TBX22* a permis de décrire un nouveau signe clinique, sous la forme d'excroissances osseuses situées à la partie postérieure du palais dur, de chaque côté de la ligne médiane palatine; et d'élargir le spectre clinique associé avec description d'atrésie choanale et séquence de Pierre-Robin. Nous avons également trouvé des mutations dans un nouveau gène, *LRRC1*, chez 2 familles distinctes. De plus nous avons étudié par WES 8 cas sporadiques ou familiaux porteurs d'un phénotype reconnaissable nommé «FLP discontinue», défini par l'association d'une fente du palais primaire et du palais secondaire avec préservation d'une partie du palais intermaxillaire. Trois mutations (>1/3) ont été trouvées dans des gènes précédemment impliqués dans les FLPs. Notre travail illustre qu'un phénotypage précis combiné à des méthodes d'investigations génétiques à haut-débit, contribuent de manière significative à une meilleure compréhension des facteurs génétiques des FLPs, ainsi qu'à la recherche clinique, épidémiologique et fondamentale, dans le but ultime d'améliorer le diagnostic et la prise en charge des patients.

KEY WORDS: WES, cleft lip and/or palate, CLP, discontinuous CLP, phenotyping, *TBX22, LRRC1*
MOTS CLES: WES, fente labio et/ou palatine, FLP, FLP discontinue, phénotypage, *TBX22, LRRC1*