



OPEN ACCESS



Open Access
Scan to access more
free content

ORIGINAL ARTICLE

A targeted next-generation sequencing assay for the molecular diagnosis of genetic disorders with orodental involvement

Megana K Prasad,¹ Véronique Geoffroy,¹ Serge Vicaire,² Bernard Jost,² Michael Dumas,² Stéphanie Le Gras,² Marzena Switala,^{3,4} Barbara Gasse,⁵ Virginie Laugel-Haushalter,⁶ Marie Paschaki,^{1,6} Bruno Leheup,⁷ Dominique Droz,⁸ Amelie Dalstein,⁸ Adeline Loing,³ Bruno Grollemund,⁴ Michèle Muller-Bolla,^{9,10} Séréna Lopez-Cazaux,¹¹ Maryline Minoux,^{3,4} Sophie Jung,^{3,4} Frédéric Oby,^{3,4} Vincent Vogt,^{3,4} Jean-Luc Davideau,⁴ Tiphaine Davit-Beal,^{5,12} Anne-Sophie Kaiser,¹³ Ute Moog,¹³ Béatrice Richard,¹⁴ Jean-Jacques Morrier,¹⁴ Jean-Pierre Duprez,¹⁴ Sylvie Odent,¹⁵ Isabelle Bailleul-Forestier,¹⁶ Monique Marie Rousset,¹⁷ Laure Merametdijan,¹⁸ Annick Toutain,¹⁹ Clara Joseph,²⁰ Fabienne Giuliano,²¹ Jean-Christophe Dahlet,³ Aymeric Courval,²² Mustapha El Alloussi,²³ Samir Laouina,²³ Sylvie Soskin,²⁴ Nathalie Guffon,²⁵ Anne Dieux,²⁶ Bérénice Doray,²⁷ Stephanie Feierabend,²⁸ Emmanuelle Ginglinger,²⁹ Benjamin Fournier,^{30,31} Muriel de la Dure Molla,^{30,31} Yves Alembik,²⁷ Corinne Tardieu,³² François Claus,^{3,4} Ariane Berdal,^{30,31} Corinne Stoetzel,¹ Marie Cécile Manière,^{3,4} Hélène Dollfus,^{1,33} Agnès Bloch-Zupan^{3,4,6}

ABSTRACT

Background Oro dental diseases include several clinically and genetically heterogeneous disorders that can present in isolation or as part of a genetic syndrome. Due to the vast number of genes implicated in these disorders, establishing a molecular diagnosis can be challenging. We aimed to develop a targeted next-generation sequencing (NGS) assay to diagnose mutations and potentially identify novel genes mutated in this group of disorders.

Methods We designed an NGS gene panel that targets 585 known and candidate genes in orodental disease. We screened a cohort of 101 unrelated patients without a molecular diagnosis referred to the Reference Centre for Oro-Dental Manifestations of Rare Diseases, Strasbourg, France, for a variety of orodental disorders including isolated and syndromic amelogenesis imperfecta (AI), isolated and syndromic selective tooth agenesis (STHAG), isolated and syndromic dentinogenesis imperfecta, isolated dentin dysplasia, otodontal dysplasia and primary failure of tooth eruption.

Results We discovered 21 novel pathogenic variants and identified the causative mutation in 39 unrelated patients in known genes (overall diagnostic rate: 39%). Among the largest subcohorts of patients with isolated AI (50 unrelated patients) and isolated STHAG (21 unrelated patients), we had a definitive diagnosis in 14 (27%) and 15 cases (71%), respectively. Surprisingly, *COL17A1* mutations accounted for the majority of autosomal-dominant AI cases.

Conclusions We have developed a novel targeted NGS assay for the efficient molecular diagnosis of a wide

variety of orodental diseases. Furthermore, our panel will contribute to better understanding the contribution of these genes to orodental disease.

Trial registration numbers NCT01746121 and NCT02397824.

INTRODUCTION

Orodental disorders encompass a number and variety of diseases that affect the teeth and oral cavity. Broadly, these disorders can be classified into anomalies of tooth number, shape and size (eg, hypo/oligo/ano-dontia (collectively selective tooth agenesis (STHAG)), microdontia, globodontia), anomalies of tooth structure (eg, amelogenesis imperfecta (AI), hereditary dentin disorders) and anomalies of tooth eruption. The prevalence of these disorders varies from relatively common (4.2% for hypodontia in the Caucasian population)¹ to extremely rare (1 in 100 000 for dentin dysplasia (DD) type 1).²

Orodental disorders can have a genetic, environmental or multifactorial basis.^{3–4} Although evidence demonstrates a role for environmental pollutants such as dioxins and fluoride in developmental enamel defects,^{5–6} a number of studies have also demonstrated a strong genetic aetiology for several orodental diseases (reviewed in refs. 7–9). Among the >5000 known genetic syndromes, >900 have orodental/craniofacial features.¹⁰ Even in the case of isolated orodental diseases, significant genetic heterogeneity exists, with several of the same genes being involved in isolated and syndromic forms of

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2015-103302>).

For numbered affiliations see end of article.

Correspondence to

Professor Agnès Bloch-Zupan, Faculty of Dentistry, University of Strasbourg, 8 rue St Elisabeth, Strasbourg 67000, France; agnes.bloch-zupan@unistra.fr

Received 6 June 2015
Revised 8 September 2015
Accepted 24 September 2015
Published Online First
26 October 2015



CrossMark

To cite: Prasad MK, Geoffroy V, Vicaire S, et al. *J Med Genet* 2016;**53**: 98–110.

disease. For instance, mutations in eight genes have been implicated in STHAG (*PAX9*, *MSX1*, *LTBP3*, *AXIN2*, *WNT10A*, *EDA*, *EDARADD* and *EDAR*),^{11–17} of which several (*MSX1*, *WNT10A*, *EDA*, *EDARADD* and *EDAR*) have also been linked with forms of ectodermal dysplasia.^{18–20} Similarly, mutations in a number of genes have been implicated thus far in AI, of which 10 cause an exclusively dental phenotype (*ENAM*, *WDR72*, *KLK4*, *AMELX*, *MMP20*, *FAM83H*, *AMBN*, *ITGB6*, *SLC24A4* and *c4orf26*),^{21–30} some cause syndromic disease with AI (*LTBP3*, *FAM20A*, *CNNM4*, *ROGDI*, *STIM1* and *FAM20C*),^{31–36} and yet others account for isolated and syndromic AI (*COL17A1*, *LAMA3*, *LAMB3* and *DLX3*).^{37–40} Indeed, the pattern of inheritance and penetrance associated with each gene also varies. This wide range of heterogeneity can render genetic diagnosis challenging.

Yet, the early molecular diagnosis of orodental disorders is important as it can improve patient care. For instance, mutations in *AXIN2* that cause STHAG have been shown to predispose carriers to colorectal cancer.¹² Early diagnosis of *AXIN2* mutations can hence alert clinicians to counsel patients to have regular colonoscopies. Similarly, the identification of mutations in *FAM20A* in patients presenting with AI can prompt a renal investigation for the management of nephrocalcinosis.⁴¹

Targeted next-generation sequencing (NGS) has proven extremely beneficial clinically for the molecular diagnosis of a number of genetically heterogeneous disorders, such as hearing loss, mitochondrial disease, intellectual disability (ID), neuromuscular disorders and Bardet–Biedl syndrome.^{42–46} Better coverage, lower cost and relative ease of data interpretation have made it more commonplace for routine clinical use than whole-exome sequencing (WES) and whole-genome sequencing (WGS).⁴⁷

We have developed the first targeted NGS panel for the molecular diagnosis of isolated and syndromic orodental disorders. We demonstrate the utility of this panel in the molecular diagnosis of a variety of orodental disorders. In a cohort of 103 patients (101 unrelated) without a known molecular diagnosis referred to the Reference Centre for Rare Diseases with Oro-dental Manifestations, we were able to provide a definitive molecular diagnosis in 39% of patients in known disease-causing genes. The identification of mutations in genes underlying syndromic forms of orodental disease highlights the potential benefits of a complete oral investigation in the diagnosis of rare genetic diseases. The aim of this article is to demonstrate the utility of this targeted NGS assay for the diagnosis of mutations in known orodental disease genes. However, the potential for the discovery of novel genes is addressed in the discussion.

METHODS

Patients

Patients were referred to the Reference Centre for Rare Diseases with Oro-dental Manifestations (Strasbourg, France) by dentists, paediatricians and geneticists from 13 hospitals and 32 private practices in France, Germany and Morocco. The inclusion criterion for the study was the presence of an orodental anomaly, defined as an anomaly of the mouth, including teeth and surrounding structures such as the periodontium (alveolar bone, ligament and gingivae), as well as defects of lip and palate formation. Patients with known mutations used for the validation assay were previously diagnosed either in clinical (Laboratory for Genetic Diagnosis, Strasbourg University Hospital) or research laboratories. DNA was obtained from peripheral blood or saliva samples (Oragene DNA, DNA Genotek, Canada).

Patient phenotype was recorded using D[4]/Phenodent (<http://www.phenodent.org>).

Gene selection and targeted capture design

Genes were selected based on their involvement in human diseases with orodental phenotypes, mutation in animal models presenting orodental disorders,^{48–49} expression in the developing mouse tooth⁵⁰ and known role in tooth development. Two versions of the gene panel were developed, v1.0 and v2.0. v1.0 was used for patients V1.01–V1.16, whereas v2.0 was used for patients V2.01–V2.95. Complementary RNA capture probes were designed against all coding exons and 25 bp of flanking intronic sequence in order to cover splice junctions of these genes using the SureDesign portal (<https://erray.chem.agilent.com/suredesign>, Agilent, USA).

Library preparation, sequencing and data analysis

Targeted regions were captured using a Custom SureSelectXT2 in-solution target enrichment kit (Agilent) and libraries were prepared for sequencing (2×100 bp) on the HiSeq2500 (Illumina, USA) following the manufacturer's instructions. For v1.0, 16 samples were multiplexed per lane for sequencing, whereas for v2.0, 32 samples were multiplexed per flow-cell lane. Read alignment, and variant calling and annotation were performed using standard methods. Briefly, reads were aligned to the GRCh37 reference genome using Burrow–Wheeler aligner (v0.7.5a)⁵¹ ensuring tagging of multi-mapped reads, and duplicates were marked with Picardv1.102 (<http://picard.sourceforge.net>). Indel realignment, base quality score recalibration and variant calling were performed with the GATK Toolkit v3.1 using hard-filtering parameters.^{52–54} Variants were annotated using snpEffv.3.4.⁵⁵ Variant frequencies were compared with an internal exome database and prioritised using VARank.⁵⁶ Variants were prioritised by allele frequency (<1% in the Single Nucleotide Polymorphism database (dbSNP137), 1000 Genomes database,⁵⁷ Exome Variant Server (EVS) database⁵⁸ and our internal database, except for non-syndromic STHAG, for which we used a cut-off of <4%) and predicted functional effect (frameshift, invariant splice sites, non-synonymous and splice affecting mutations). Missense mutations were evaluated for pathogenicity bioinformatically using Sorting Intolerant from Tolerant (SIFT), PolyPhen, MutationTaster and amino acid conservation.^{59–61} Splice affecting mutations were evaluated bioinformatically using Human Splice Finder, MaxEntScan, NNSplice, Gene Splicer, SSF, Rescue ESE and ESE Finder.^{62–67} CNVs were detected as previously described.^{44–45} Variant pathogenicity was interpreted according to the American College of Medical Genetics guidelines.⁶⁸ Evidence used to establish pathogenicity is provided in online supplementary figures S1–S45.

Mutation validation

All mutations were validated by Sanger sequencing (GATC, Germany). Segregation analyses were performed whenever DNA was available for additional family members. Single-molecule PCR followed by Sanger sequencing was used to phase biallelic mutations when parental DNA was unavailable.^{69–70} Deletions were validated by qPCR. The region of interest and an internal control region (*RPPH1*) were amplified from 2 ng of genomic DNA from the patient and a control using the iQ SYBR Green Supermix (170-8880, Biorad, USA) on a CFX96 Real-Time System (Biorad). Data were analysed using the Pfaffl method.⁷¹

RESULTS

Validation of v1.0 of the NGS panel

As a proof of principle, we created a primary version (v1.0) of a custom NGS panel that targets 560 known and candidate genes (2.36 Mb) in orodental disorders. Of these genes, probes for 175 known and strong candidate genes (0.81 Mb) were designed to have superior coverage (diagnosis subpanel), whereas the remaining candidate genes constituted a 'Discovery subpanel' (see online supplementary tables S1 and S2). We validated v1.0 in a cohort of 16 patients: patients V1.01–V1.08 with isolated and/or syndromic orodental diseases with known mutations as determined by previous candidate gene Sanger sequencing, WES, or array CGH (see online supplementary table S3) and patients V1.09–V1.16 without a known molecular diagnosis. Patients V1.01–V1.08 were selected to include a variety of heterozygous, homozygous and hemizygous mutations (substitutions, indels, a large heterozygous deletion) at eight different loci in order to test the sensitivity of the gene panel to different types of genetic alterations. Using our computational pipeline, 10 of 10 mutations, including a large heterozygous deletion on the X chromosome (see online supplementary table S3), were detected in a blinded manner by a bioinformatician who was unaware of the molecular alteration but was aware of the clinical diagnosis for each patient. The bioinformatician was also blinded to the mode of inheritance in order to simulate a diagnostic scenario where such information is often unavailable. Samples V1.09–V1.16 were selected to cover a variety of orodental phenotypes (AI, dentinogenesis imperfecta (DGI), DD, STHAG) to determine the potential to identify unknown mutations with our panel. The molecular results from samples V1.09–V1.16 are presented and discussed in tables 1–3, together with the results from v2.0 of the gene panel presented below. Briefly, pathogenic or likely pathogenic mutations were identified in six of eight samples among samples V1.09–V1.16 (tables 2 and 3).

On average, in this validation cohort, by multiplexing 16 samples per lane of a sequencing flow cell, 2.2 Gb of sequence was generated per sample, giving a mean coverage of 365 \times , with 95.31% nucleotides covered at $\geq 50\times$ (see online supplementary table S4). The mean coverage of the diagnostic subpanel was 404 \times , with 98.94% of nucleotides covered at $\geq 20\times$, permitting confident diagnosis of mutations in these genes, for which we set a minimum coverage threshold of 20 \times (see online supplementary table S4). A small percentage of targeted regions (33.3 kb, 1.4%) had an average coverage $< 20\times$ across samples due to high guanine-cytosine (GC) content or pseudogenes with highly similar sequences (see online supplementary table S5). The majority of these regions were in the discovery subpanel (28.1 kb). Under-representation of high GC content and pseudogenic regions is a known issue with probe-based targeting strategies. However, these gaps can be filled in by Sanger sequencing for regions considered to be relevant on a case-by-case basis. On average, we detected ~2648 variants/sample, of which ~87 variants/sample were rare ($< 1\%$ in dbSNP137, 1000Genomes, EVS and an in-house database). Among these rare variants, we identified on average per sample ~20 missense changes, 0.38 nonsense mutations, 0.25 splice-site changes and 0.56 frameshift-inducing indels (see online supplementary table S6).

Development of v2.0 of the NGS panel

Having validated v1.0 of our NGS gene panel, we included 25 additional genes to the discovery subpanel to create a second

version of the panel (v2.0) that included 585 genes (2.47 Mb). These additional genes were included due to their implication in animal models of orodental disease that was unknown when designing v1.0 of the panel v2.0 of the panel was used to sequence 95 patients (V2.01–V2.95). Furthermore, since the average coverage achieved with v1.0 was more than sufficient for confident molecular diagnosis, with v2.0, 32 samples were multiplexed per well of a flow cell lane in order to reduce sequencing costs while ensuring a minimum average coverage of 100 \times per sample. Sequencing output with v2.0 is shown in online supplementary table S7. Briefly, we achieved a mean coverage of 179 \times overall, with 97.2% of the targeted region covered at $\geq 20\times$. Furthermore, the diagnostic panel had an average coverage of 211 \times with 97% of bases covered at $\geq 50\times$.

Screening a cohort of patients with diseases with orodental involvement

A description of the final cohort of 103 patients (101 unrelated) without a known mutation sequenced with v1.0 (V1.09–V1.16) and v2.0 (V2.01–2.95) and the diagnostic yield by disease category is shown in table 1. Isolated AI was the most common disorder in the cohort (50%), followed by isolated STHAG (20%). We also included several patients with syndromic forms of AI, ranging from well-defined syndromes, such as Enamel Renal Syndrome, to undefined and suspected syndromes, hoping that a molecular diagnosis may aid difficult clinical diagnosis. The remaining cases consisted of patients with isolated or syndromic dentin disorders, syndromic STHAG, suspected otodental dysplasia and primary failure of tooth eruption. Since AI and STHAG can be inherited in an autosomal-dominant (AD),

Table 1 Cohort description and diagnostic yield per disease category

Disease	Number of patients	Diagnostic yield, N (%)
Isolated AI	52 (51*)	14 (27%)
Of which confirmed AI	51 (50*)	15 (29%, 30%*)
Of which suspected AI	1	0 (0%)
Syndromic AI	14	1 (7%)
Enamel renal syndrome	1	1 (100%)
Mucopolysaccharidosis IV A	1	0 (0%)†
Kohlschutter Tonz (suspected)	1	0 (0%)
Osteogenesis imperfecta (suspected)	1	0 (0%)
Spondyloepiphyseal dysplasia	1	0 (0%)†
Undefined syndrome	9	0 (0%)
Isolated STHAG	21	15 (71%)
Syndromic STHAG	4	2 (50%)
Ectodermal dysplasia	3	1 (33%)
Intellectual disability with STHAG	1	1 (100%)
Isolated DGI	5 (4*)	5, 4* (100%)
Syndromic DGI	2	1 (50%)
Goldblatt syndrome (suspected)	1	0 (0%)
Osteogenesis imperfecta	1	1 (100%)
Isolated DD	2	1 (50%)
Otodental syndrome	1	1 (100%)
Primary failure of tooth eruption	2	0 (0%)
Total	103 (101*)	40, 39* (39%)

*Number of unrelated patients.

†Likely pathogenic mutations were identified in *GALNS* in these patients.

AI, amelogenesis imperfecta; DD, dentin dysplasia; DGI, dentinogenesis imperfecta; STHAG, selective tooth agenesis.

Table 2 Pathogenic mutations identified in patients without a known mutation

Patient ID	Sex	Age (years)	Mode	Clinical features	Gene	Transcript	c.	p.	Inheritance	Reference	Figure
<i>Pathogenic variants</i>											
<i>Isolated AI</i>											
V1.14	M	25	?	Hypoplastic, hypomineralised	<i>WDR72</i>	NM_182758.3	c.[182A>G];[815G>A]	p.[H61R];[W272*]	Compound heterozygous	Novel	S1
V2.05	M	6	AR	N/A	<i>CNNM4</i>	NM_020184.3	c.[1495G>A];[1495G>A]	p.[V499M];[V499M]	Homozygous (C)	Novel	S2
V2.08	M	12	AR	Hypomineralised	<i>SLC24A4</i>	NM_153646.3	c.[(1537+1_1538-1)_(*67_?)del]; [(1537+1_1538-1)_(*67_?)del]	p.[0?];[0?]	Homozygous (C)	72	S3
V2.09	F	16	AD	Hypoplastic	<i>COL17A1</i>	NM_000494.3	c.[2407G>T]; [=]	p.[G803*];[=]	Maternal	73	S4
V2.18	M	13	?	Hypoplastic	<i>LAMB3</i>	NM_000228.2	c.[124C>T]; [=]	p.[R42*];[=]	?	74	S5
V2.20	F	18	?	Hypomature	<i>FAM83H</i>	NM_198488.3	c.[1289C>A];[=]	p.[S430*];[=]	?	75	S6
V2.26	M	13	AD	Hypomineralised	<i>FAM83H</i>	NM_198488.3	c.[1282C>T];[=]	p.[Q428*];[=]	Paternal	Novel	S7
V2.28	M	14	?	Hypoplastic	<i>AMELX</i>	NM_182680.1	c.[155C>T];[=]	p.[P52L];[=]	De novo	Novel	S8
V2.29	M	14	AD	N/A	<i>ENAM</i>	NM_031889.2	c.[123+1G>A];[=]	p.[0?];[=]	Paternal	Novel	S9
V2.48	F	20	AD	Hypoplastic	<i>COL17A1</i>	NM_000494.3	c.[1646G>A];[=]	p.[W549*];[=]	Maternal	Novel	S10
V2.53	F	19	?	Hypoplastic	<i>AMBN</i>	NM_016519.5	c.[532-1G>C];[532-1G>C]	p.[0?];[0?]	Homozygous	Novel	S11
V2.63	M	10	?	N/A	<i>FAM83H</i>	NM_198488.3	c.[2029C>T];[=]	p.[Q677*];[=]	?	76	S12
V2.79	M	10	?	Hypoplastic, hypomature	<i>COL17A1</i>	NM_000494.3	c.[1873C>T];[=]	p.[R625*];[=]	?	Novel	S13
V2.82	F	14	AD	Hypoplastic	<i>COL17A1</i> ; <i>LAMA3</i>	NM_000494.3; NM_198129.1	<i>COL17A1</i> :c.[1141+1G>A]; <i>LAMA3</i> :c.[6477_6486del]	<i>COL17A1</i> :p.[0?]; <i>LAMA3</i> :p.[I2159Mfs*46]	Maternal and ?	Novel	2
<i>Syndromic AI</i>											
V2.06	F	14	?	Enamel Renal Syndrome	<i>FAM20A</i>	NM_017565.3	c.[1106_1107del]; [c.1006_1107del]	p.[E2316Gfs*10]; [E2316Gfs*10]	Homozygous	Novel	S14
<i>Isolated STHAG</i>											
V1.11	M	21	?	Ag 12, 13, 14, 15, 18, 22, 23, 24, 25, 28, 31, 34, 35, 38, 44, 45, 48	<i>WNT10A</i>	NM_025216.2	c.[383G>A];[=]	p.[R128Q];[=]	Maternal	77	S15
V1.15	F	11	?	Ag 12, 15, 17, 22, 25, 27, 28, 31, 35, 37, 38, 41, 44, 45, 47, 48	<i>WNT10A</i>	NM_025216.2	c.[343A>C](c);[682T>A]	p.[K115Q](c);[F228I]	?	Novel and 78	S16
V2.55	F	15	?	Ag 11, 12, 13, 14, 17, 18, 21, 22, 23, 24, 25, 27, 28, 31, 32, 33, 34, 37, 38, 41, 42, 43, 44, 45, 47, 48	<i>WNT10A</i>	NM_025216.2	c.[321C>A];[321C>A]	p.[C107*];[C107*]	Homozygous	78	S17
V2.65	F	17	?	Ag 12, 13, 15, 17, 18, 22, 23, 25, 27, 28, 31, 32, 35, 37, 38, 41, 42, 45, 48	<i>WNT10A</i>	NM_025216.2	c.[682T>A];[682T>A]	p.[F228I];[F228I]	Homozygous	78	S18
V2.66	M	37	?	Ag 12, 13, 18, 22, 28, 31, 32, 37, 38, 41, 42, 47, 48	<i>WNT10A</i>	NM_025216.2	c.[682T>A];[682T>A]	p.[F228I];[F228I]	Homozygous	78	S19
V2.67	M	11	AD	Ag 15, 16, 17, 18, 25, 26, 27, 28, 36, 37, 38, 45, 46, 47, 48	<i>PAX9</i>	NM_006194.3	c.[(?_-115)_(*62_?)del];[=]	p.[0?];[=]	Paternal	78	S20
V2.69	M	18	?	Ag 13, 14, 18, 22, 23, 24, 28, 31, 41, 45	<i>WNT10A</i>	NM_025216.2	c.[637G>A];[=]	p.[G213S];[=]	?	79	S21
V2.71	M	11	?	Ag 12, 13, 15, 17, 18, 22, 23, 25, 27, 28, 31, 32, 33, 34, 35, 37, 38, 41, 42, 43, 44, 45, 47, 48	<i>WNT10A</i>	NM_025216.2	c.[682T>A];[321C>A]	p.[F228I];[C107*]	Compound heterozygous	78	S22
V2.72	M	33	?	Ag 17, 18, 22, 28, 31, 32, 37, 38, 41, 42, 47, 48	<i>WNT10A</i>	NM_025216.2	c.[682T>A](c);[321C>A]	p.[F228I](c);[C107*]	?	78	S23
V2.74	M	26	?	Ag 13, 15, 18, 23, 24, 25, 28, 31, 33, 34, 35, 37, 38, 41, 42, 43, 44, 45, 47, 48	<i>WNT10A</i>	NM_025216.2	c.[682T>A];[682T>A]	p.[F228I];[F228I]	Homozygous	78	S24

Continued

Table 2 Continued

Patient ID	Sex	Age (years)	Mode	Clinical features	Gene	Transcript	c.	p.	Inheritance	Reference	Figure
V2.76	M	15	?	Ag 12, 14, 15, 18, 22, 24, 25, 28, 31, 34, 35, 37, 38, 41, 42, 43, 44, 45, 47, 48	<i>WNT10A</i>	NM_025216.2	c.[682T>A];[682T>A]	p.[F228I];[F228I]	Homozygous	78	S25
V2.78	F	9	?	Ag 12, 14, 22, 31, 32, 35, 41	<i>EDA</i>	NM_001399.4	c.[467G>A];[=]	p.[R156H];[=]	De novo	17	S26
V2.91	F	12	AD	Ag 14, 15, 24, 25, 35, 36, 45, 46	<i>MSX1</i>	NM_002448.3	c.[249del];[=]	p.[E84Rfs*76];[=]	Paternal	Novel	S27
V2.92	M	9	?	Ag 12, 17, 22, 23, 24, 25, 27, 35, 37, 33, 31, 41, 43, 45, 47	<i>WNT10A</i>	NM_025216.2	c.[682T>A];[=]	p.[F228I];[=]	Maternal	78	S28
V2.93	M	28	?	Ag 14, 15, 18, 25, 27, 28, 32, 34, 35, 37, 38, 44, 45, 47, 48	<i>WNT10A</i>	NM_025216.2	c.[682T>A];[682T>A]	p.[F228I];[F228I]	Homozygous	78	S29
Syndromic STHAG											
V2.54	F	49	?	ED	<i>WNT10A</i>	NM_025216.2	c.[682T>A];[c.416C>T]	p.[F228I];[p.A139V]	Compound heterozygous	78 and novel	S30
V2.87	F	3	?	Intellectual disability	<i>CTNNB1</i>	NM_001904.3	c.[998dup];[=]	p.[Tyr333*];[=]	?	Novel	S31
Isolated DGI/DD											
V1.09	F	10	AD	DD*	<i>DSPP</i>	NM_014208.3	c.[3480del];[=]	p.[S1160Rfs*154];[=]	?	Novel	S32
V1.10	M	9	AD	DGI*	<i>DSPP</i>	NM_014208.3	c.[3480del];[=]	p.[S1160Rfs*154];[=]	Maternal	Novel	S32
V2.36	M	45	AD	DGI	<i>DSPP</i>	NM_014208.3	c.[3533_3534insTA];[=]	p.[N1179Tfs*136];[=]	?	Novel	S33
V2.55	F	12	AD	DGI	<i>DSPP</i>	NM_014208.3	c.[52G>T];[=]	p.[V18F];[=]	Maternal	80	S34
V2.57	F	44	AD	DGI	<i>DSPP</i>	NM_014208.3	c.[3480del];[=]	p.[S1160Rfs*154];[=]	?	Novel	S35
V2.59	F	41	?	DGI	<i>DSPP</i>	NM_014208.3	c.[3682_3686del];[=]	p.[S1228*];[=]	?	Novel	S36
Syndromic DGI											
V2.58	M	47	AD	Osteogenesis imperfecta	<i>COL1A1</i>	NM_000088.3	c.[3837_3840del];[=]	p.[N1279Lfs*51];[=]	Paternal	Novel	S37
Otodontal syndrome											
V2.03	M	13	AD		<i>FGF3</i>	NM_005247.2	c.[(?_-3)_(220+1_221-1)del];[=]	p.[0?];[=]	Paternal	Novel	S38

*Related individuals.

AD, autosomal dominant; AI, amelogenesis imperfecta; AR, autosomal recessive; C, consanguineous parents; DD, dentin dysplasia; DGI, dentinogenesis imperfecta; ED, ectodermal dysplasia; ID, intellectual disability; STHAG, selective tooth agenesis.

Table 3 Likely pathogenic variants and other variants identified in patients without a known mutation

Patient ID	Sex	Age (y)	Mode	Clinical features	Gene	Transcript	c.	p.	Inheritance	Reference	Figure
<i>Likely pathogenic variants</i>											
V1.12	M	11	?	Ectodermal dysplasia	EDA	NM_001399.4	c.[396+5G>A];[=]	p.[0?];[=]	Maternal	Novel	S39
V2.07	M	14	?	AI hypomineralised, hypoplastic	MMP20	NM_004771.3	c.[954-2A>T];[126+6T>G]	p.[0?];[0?]	Compound heterozygous	²⁵ and novel	S40
V2.13	F	17	?	AI hypoplastic, hypomature	MMP20	NM_004771.3	c.[954-2A>T](c);[c.103A>C]	p.[0?](c)[R35R]	?	²⁵ and novel	S41
V2.15	M	20	?	Spondyloepiphyseal dysplasia	GALNS	NM_000512.4	c.[121-31];[935C>G]	p.[0?];[T312S]	Compound heterozygous	Novel and ⁸¹	S42
V2.32	F	26	?	AI hypomineralised, hypoplastic	FAM20A	NM_017565.3	c.[590-2A>G;590-3C>A](c);[1294G>A]	p.[0?](c)[A432T]	Maternal and de novo	⁸² and novel	S43
V2.86	F	6	?	AI hypoplastic	LAMB3	NM_000228.2	c.[1903C>T];[=]	p.[R635*];[=]	Maternal, asymptomatic	⁸³	S44
Cases where only a single pathogenic variant was identified											
V2.49	F	4	?	Mucopolysaccharidosis IVA	GALNS	NM_000512.4	c.[1156C>T];[?]	p.[R386C];[?]	de novo	⁸⁴	S45

AI, amelogenesis imperfecta; ID, intellectual disability.

autosomal-recessive (AR) or X-linked fashion, the mode of transmission was unclear for the majority of cases. Sixty-six cases showed sporadic disease, whereas 35 showed familial transmission. The mode of inheritance could be confidently inferred only in 18 cases, 15 as AD due to the presence of other affected family members, 2 as AR due to the presence of consanguinity between the parents of the index case and 1 as X-linked. Due to the large number of cases without a known mode of inheritance of disease, a common situation in diagnostic cohorts, variant filtration was performed for all samples for all modes of inheritance. Overall, we had a definitive molecular diagnosis in known genes in 39 cases (39%) and discovered 21 novel pathogenic variants. All pathogenic mutations identified in this cohort are listed in table 2. Likely pathogenic mutations and other variants are shown in table 3. The distribution of pathogenic variants identified with respect to the disease category and the number of cases corresponding to each mutated gene is shown in figure 1.

Isolated AI: *COL17A1* mutations are a frequent cause of AI. Of the 50 unrelated cases with confirmed isolated AI, we identified the causative mutation in 14 cases (27%). Surprisingly, the most frequently mutated gene in our cohort was *COL17A1*, with four independent patients segregating pathogenic mutations in this gene (of which three are novel). In the biallelic state, mutations in *COL17A1*, encoding a structural component of hemidesmosomes, cause junctional, non-Herlitz-type epidermolysis bullosa (EB), a skin blistering disorder with associated dental enamel defects.⁸⁵ However, rare cases of isolated ADAI in heterozygous carriers have been reported.^{37 86} Similarly, heterozygous mutations in *LAMA3* and *LAMB3*, which also encode hemidesmosomal components, can cause ADAI, whereas biallelic mutations in these genes cause EB.^{38 39} We found one family segregating a known heterozygous *LAMB3* nonsense mutation. These patients showed a similar enamel phenotype characterised by a hypoplastic form of AI with thin enamel and a characteristic pitting of the enamel that varied in severity (figures 2A–H and 3C). Depending on the extent of pitting, the enamel surface was sometimes rough. Secondary extrinsic colouration of the pits rendered them more visible. Interestingly, patient V2.82 had two unlinked mutations, one maternally inherited novel *COL17A1* splice mutation (c.1141+1G>A) and a novel *LAMA3* frameshift-inducing deletion (p.I2159Mfs*46) that is absent in her mother (figure 2I). The father of the patient was unavailable for testing but was described as unaffected at the time of ascertainment. The c.1141+1G>A mutation is predicted to cause either in-frame skipping of exon 14 (which encodes a part of the intracellular domain) or retention of intron 14 that would subsequently introduce a premature termination codon (PTC) and likely nonsense-mediated decay (NMD) of the transcript. Patient V2.82's enamel phenotype is more severe than that of her mother, suggesting digenic inheritance in this individual. Patient V2.82 has hypoplastic AI with clearly visible pitted enamel in the vestibular, lingual and palatal surfaces of premolars and molars and a white colouration following Retzius striae visible in the permanent incisors (figure 2A, B). Her mother's dentition, in contrast, shows more discrete signs of AI and fewer pits are visible on the buccal surfaces of premolars (figure 2C, D). Digenic mutations in *LAMB3* and *COL17A1* that modify the severity of the EB phenotype have been previously reported.⁸⁷ However, this is the first report of seemingly digenic inheritance in ADAI.

Figure 1 Distribution of pathogenic variants identified by disease category. The genes in which pathogenic mutations were identified are shown with respect to the disease in which they were identified. In parentheses next to each gene is the number of independent patients in whom mutations were identified in the gene. This figure includes results from patients V1.09–V1.16 and V2.01–V2.95. DD dentin dysplasia; DGI, dentinogenesis imperfecta; STHAG selective tooth agenesis.

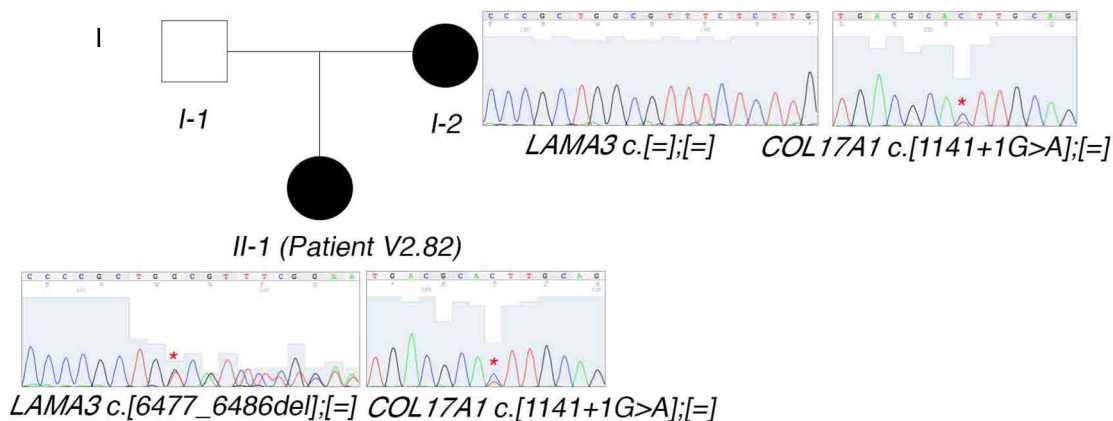
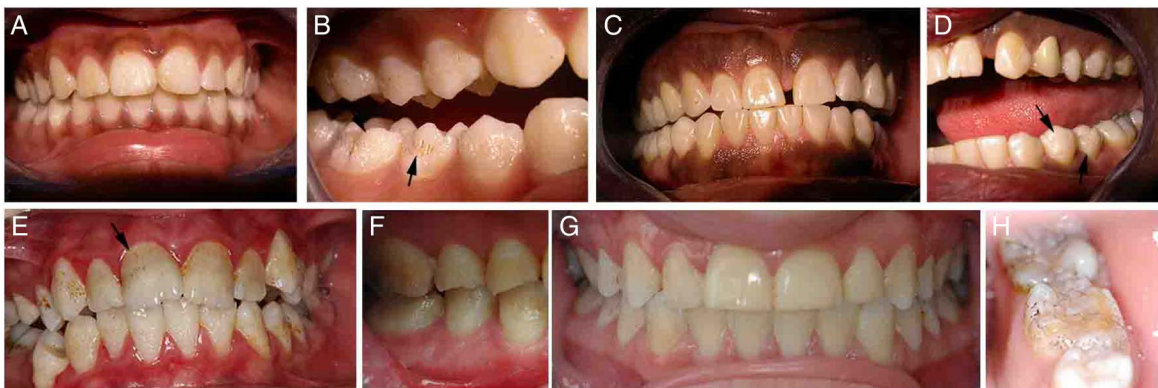
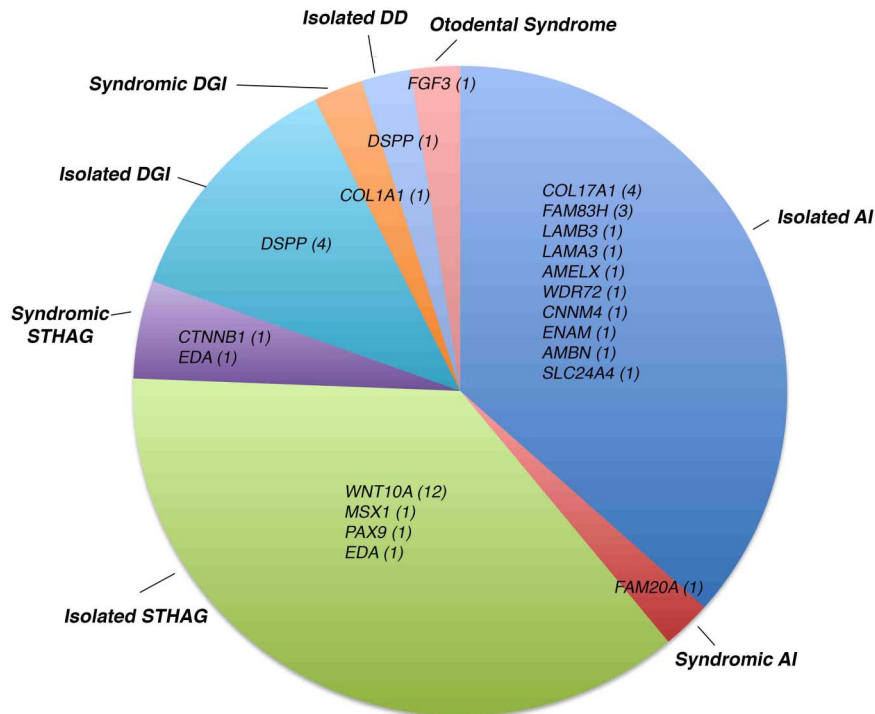


Figure 2 *COL17A1* mutations show a characteristic enamel phenotype in autosomal-dominant amelogenesis imperfecta. (A–H) Photos of the enamel phenotypes of individuals with *COL17A1* mutations. (A, B) Patient V2.82. (C, D) Mother of patient V2.82. (E, F) Patient V2.09. (G, H) Patient V2.48. Arrows mark pits in the enamel that are characteristic of *COL17A1* mutations. Extrinsic colouration makes these pits more visible in (B), (E) and (H). (I) Seemingly digenic inheritance of amelogenesis imperfecta in patient V2.82. Asterisks on the sequence chromatograms mark the mutated nucleotides.

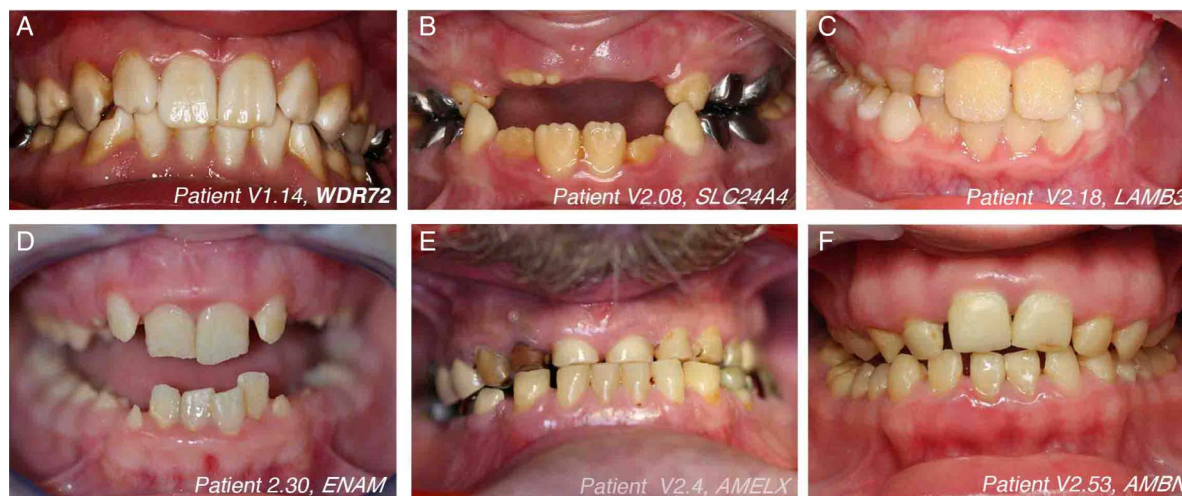


Figure 3 The variety of isolated amelogenesis imperfecta phenotypes seen in this cohort. Representative images of the enamel phenotype associated with mutations in different genes. The mutated gene and patient number are indicated in each panel.

We also identified a family with a novel homozygous splice-site mutation (c.532-1G>C) in *AMBN*. To date, only a single family has been reported to carry mutations in this gene.²⁷ Similar to the patients in the initial study, our patient too exhibited hypoplastic AI with very limited enamel as seen in the unrestored yellow-tinged premolars and permanent molars. However, the tooth surfaces did not seem as pitted as described previously (figure 3F).²⁷ This mutation is predicted to alter mRNA splicing by inducing either the skipping of exon 7 or the retention of intron 6. Exon skipping would lead to in-frame deletion of exon 7, which encodes a domain involved in heparin and fibronectin interaction that is thought to be important for the interaction of *AMBN* with dental epithelial cells.^{88 89} Intron retention would introduce a PTC that is likely to cause NMD of the mRNA. Additionally, a cryptic acceptor splice-site is predicted bioinformatically to be activated in exon 7 that would cause partial out-of-frame skipping of exon 7 and a subsequent PTC (see online supplementary figure S11). Thus, this mutation is predicted to cause at least a partial loss of *AMBN* function. We also identified novel mutations in *WDR72*, *AMELX* and *ENAM*, and a known deletion of the last three exons of *SLC24A4* (table 1). Representative images of the enamel phenotype associated with mutations in each gene are shown in figure 3. Similar to previous reports, *WDR72* and *SLC24A4* mutations caused hypomineralised, hypomature AI with brownish/yellowish discoloured teeth.^{22 29} The enamel showed a lack of contrast with the underlying dentin upon X-ray imaging (data not shown). The softer enamel seen in the patient with *WDR72* mutations was subject to wear, whereas the enamel in the patient with the *SLC24A4* mutation was opaque. Patients with *AMELX*, *LAMB3* and *ENAM* mutations had hypoplastic AI with yellow discoloration and smaller teeth due to thinner enamel.

Interestingly, we identified a novel homozygous missense mutation (p.V499M) in *CNNM4* in patient V2.05 who was born to first-cousins and was referred for isolated AI. This mutation affects a highly conserved residue in a functional cystathionine beta-synthase domain of the protein and is predicted bioinformatically by three algorithms (SIFT, PolyPhen and MutationTaster) to be deleterious (see online supplementary figure S2). However, mutations in this gene are known to cause Jalili syndrome, which is characterised by a combination of AI and cone-rod dystrophy.^{33 90} A subsequent full-field

electroretinography revealed a marked loss of cone response and a less severe loss of rod response in the patient, thus confirming the Jalili phenotype (see online supplementary figure S2E, F).

Syndromic AI

We confirmed the molecular diagnosis in one case of Enamel Renal Syndrome (patient V2.06). This patient was homozygous for a novel frameshift-inducing deletion in *FAM20A* (c.1106_1107del).

In patient V2.15, referred for the management of AI in the context of suspected spondyloepiphyseal dysplasia (SED), we identified a known pathogenic missense mutation (p.T312S) and a likely pathogenic intronic variant (c.121-31) in *GALNS* (mutated in mucopolysaccharidosis (MPS) IVA) predicted to create an intronic splice enhancer (see online supplementary figure S42). However, the biological effect of this intronic mutation warrants further investigation. No deleterious mutations were identified in *CHST3* or *COL2A1* and a differential diagnosis of mucopolysaccharidosis type III was eliminated due to normal hexosaminidase, alpha-L-fucosidase and beta-D-glucuronidase levels (data not shown). Similar clinical presentation of SED and MPS IVA has been previously demonstrated;⁹¹ therefore, our results suggest a potential differential diagnosis of MPS IVA in this patient. In the case of patient V2.49, referred for the management of AI linked to MPS IVA,^{92 93} we identified only a single de novo known pathogenic mutation in *GALNS* (p.R386C), suggesting the presence of a second mutation in intronic or regulatory regions, or a structural mutation missed by our pipeline. The presence of an undetected second mutation is consistent with the virtually absent *GALNS* activity in the patient's leucocytes (1 nmol/h/mg protein) (data not shown). Both these patients exhibited hypoplastic AI with thin striated and pitted enamel and subsequently flattened buccal surfaces (see online supplementary figure S45) consistent with previous reports of the enamel phenotype seen in MPS IVA,⁹⁴ further supporting a diagnosis of MPS IVA in both patients. However, additional tests are necessary to confirm this diagnosis.

Isolated and syndromic STHAG

We identified the causative mutation in 15/21 (71%) cases of isolated STHAG. The majority carried mutations in *WNT10A* either in a heterozygous, homozygous or compound

heterozygous state as has been previously reported.^{77–79} Fortuitously, we were also able to provide a molecular diagnosis for a patient presenting with ID who was referred for the management of hypodontia. The patient (V2.87) had agenesis of five incisors, two maxillary incisors and three mandibular permanent incisors (agenesis 12, 22, 41, 31 and 32) and had no family history of hypodontia or ID. We identified a novel heterozygous nonsense mutation in *CTNNB1* (p.Tyr333*), which encodes beta-catenin, a member of the WNT signalling pathway, which is essential during multiple stages of tooth development.⁹⁵ Recent reports have established a role for heterozygous *CTNNB1* mutations in ID characterised by mild to severe ID, autism spectrum disorder childhood hypotonia with progressive spastic diplegia, microcephaly and significant additional craniofacial and brain abnormalities, which is compatible with this patient's phenotype (data not shown).^{96–97} However, the orodental phenotype in this syndrome has not been studied in detail. Our report suggests that patients with ID due to *CTNNB1* mutations may require a dental examination and management of hypodontia. Furthermore, this case highlights the benefit of an interdisciplinary approach to patient care and the potential utility of a full dental examination in syndromes with craniofacial involvement.

Dentinogenesis imperfecta and dentin dysplasia

Among the six independent cases of isolated dentin disorders, we identified the causative mutation in five. All patients carried mutations in *DSPP*, the only gene implicated thus far in isolated DD/DGI.² The majority of mutations were present in the repeat-rich exon 5 that is often refractory to Sanger sequencing and hence until now excluded from routine diagnostic sequencing.⁹⁸ We achieved an average coverage of 771× and 408× over exon 5 of *DSPP* with v1.0 and v2.0, respectively. Hence, our assay provides an effective means to screen for mutations in the last exon of *DSPP*. Our bioinformatic pipeline detected the majority of mutations; however, one mutation (c.3682_3686del) was missed due to the difficulty of calling variants in structurally complex regions (personal communication, MKP, Broad Institute GATK Team), but was detected upon visual inspection of the reads. We also identified a novel frameshift-inducing deletion in *COL1A1* in a family segregating AD osteogenesis imperfecta with DGI.

Otodental dysplasia

We identified a heterozygous deletion of the first exon of *FGF3* in a family segregating suspected AD otodental dysplasia, which is characterised by globodontia and hearing loss, sometimes segregating with colobomas, most likely due to the deletion of the adjacent *FADD* gene.⁹⁹ This deletion was validated by qPCR on genomic DNA, which showed the presence of the deletion in patient V2.03 as well as in his affected father (see online supplementary figure S38).

DISCUSSION

We have developed a targeted NGS assay for the diagnosis and discovery of mutations underlying rare genetic disorders with orodental manifestations. We achieved high sequencing coverage in the targeted regions, 179× average coverage with 97.2% of the targeted region covered at $\geq 20\times$ with V2.0. In a cohort of 101 unrelated patients with a variety of orodental genetic disorders of unknown genetic aetiology, we were able to detect the underlying pathogenic mutation in 39 cases (39%) in known genes.

Targeted NGS gene panels are being widely used for the diagnosis of a variety of genetic disorders, including ID, mitochondrial disorders and neuromuscular diseases.^{42–44} However, to the best of our knowledge, this is the first report of a targeted NGS gene panel for orodental disorders. Therefore, this diagnostic tool responds to a veritable need within the dental genetics community. Furthermore, two concrete examples from our cohort suggest the utility of this tool to the wider medical genetics community. First, we were able to diagnose a mutation in *CTNNB1* in a patient with ID and tooth agenesis, suggesting the potential utility of our assay in screening for mutations in ID cases with orodental involvement. Second, the detection of a *CNNM4* mutation in a patient referred for isolated AI led to an ophthalmological investigation for cone-rod dystrophy and the confirmation of Jalili syndrome. These cases highlight how the oral consultation can be a port of entry for the diagnosis and management of patients with rare genetic diseases, especially in light of the involvement of the same genes in syndromic and isolated forms of orodental disease and the sometimes non-evident extra-oral signs associated with some diseases.

Our targeted NGS panel targets fewer genes than WES/WGS and may thus exclude potential disease-causing genes. Yet, it has several advantages in a clinical setting. First, it provides higher sequencing coverage in the targeted regions than WES/WGS. A comparison of the coverage of the same regions achieved with v2.0 of our gene panel and a commercially available WES kit demonstrates that the coverage from WES falls short of the requirements for confident molecular diagnosis, that is, only 89.09% of targets covered at $\geq 20\times$ with WES vs 97.23% of targets covered at $\geq 20\times$ with our panel (see online supplementary figure S46 and table S8). Second, due to the smaller number of variants identified by targeted sequencing (~2600 variants with our panel vs ~30 000 with WES), variant interpretation is relatively simpler. This is also true for CNVs—we were able to detect and validate a single exon deletion in *FGF3*. Finally, the cost, amount of time for data analysis and interpretation, and data storage requirements are lower with targeted NGS than with WES/WGS approaches, making it an attractive alternative for clinical implementation. Alternatively, our assay could serve as a primary mutation-screening tool to exclude mutations in known genes before performing WES/WGS.

We expected that the inclusion of candidate genes selected based on their expression in developing mouse teeth⁵⁰ and their implication in animal models of orodental disorders would permit the identification of novel genes mutated in these disorders. In order to identify potentially novel genes in patients lacking mutations in known genes, we performed a preliminary analysis focusing on rare loss-of-function variants (nonsense, invariant splice site and frameshift variants) that were present in the same gene in at least two independent families with the same phenotype. Such a strategy was previously used successfully to identify a novel gene, *SETD5*, in ID.¹⁰⁰ However, this preliminary analysis did not yield any positive findings. Indeed, this analysis was limited in that it did not consider missense mutations whose effect is less clear and was complicated by the lack of information regarding the mode of inheritance in the majority of cases. Also, in this analysis patients were regrouped by disease category (all mutation-negative AI patients, all mutation-negative STHAG patients) without subclassification based on fine phenotype (eg, hypoplastic vs hypomineralised AI and severity of STHAG), which may be important to include in order to overcome the difficulties posed by genetic heterogeneity.¹⁰¹ Furthermore, the rarity of mutations in some genes may also necessitate much larger cohorts. Indeed, future analyses of this cohort will try to address these limitations.

Among the 50 unrelated patients with confirmed isolated AI in our cohort, we had a molecular diagnosis rate of 27%. This is slightly lower than the diagnostic rates reported by other studies that performed candidate gene sequencing in large AI cohorts. By performing Sanger sequencing of six genes (*FAM83H*, *ENAM*, *AMELX*, *MMP20*, *KLK4* and *WDR72*) in 71 families, Wright *et al*⁷⁵ identified mutations in 26 families (37%). By sequencing the same six genes in an independent cohort of 39 AI kindreds, Chan *et al*¹⁰² identified mutations in 19 kindreds (49%). In the latter study, the diagnostic rate was shown to be dependent on the mode of transmission: X-linked>AD>AR>simplex cases. In our study, the majority of cases were simplex cases (31 cases), which may explain to a degree our lower diagnostic rate despite the inclusion of a larger number of known AI genes. Furthermore, the 50 patients selected for this study belong to a larger AI cohort of 71 patients, a part of which (including patients in this study) had been previously sequenced by candidate gene Sanger sequencing and in which mutations were identified in 11 patients¹⁰³ (online supplementary table S3 and unpublished data). Therefore, the overall diagnostic rate in the 71-patient cohort is 35% (25 patients), which is closer to published diagnostic rates. In further contrast to the above-mentioned studies where *FAM83H* was the most frequently mutated gene in ADAI, *COL17A1* was the most frequently mutated gene in ADAI in our cohort, with 8% of our AI cohort carrying deleterious mutations in *COL17A1*. Additionally, we demonstrate that the presence of unlinked mutations in *COL17A1* and *LAMA3* can modify the severity of AI, suggesting a digenic mode of AI inheritance associated with mutations in the hemidesmosomal components. Therefore, our findings are novel in demonstrating the frequency of *COL17A1* mutations in ADAI. Finally, the large number of AI patients without mutations in known genes suggests that the genetic and allelic heterogeneity underlying AI is yet to be fully unravelled.

The diagnostic rate in syndromic AI was low in our cohort (7%). This is likely because of the inclusion of patients without a clear clinical diagnosis. These cases likely represent novel phenotypes with mutations in novel genes. Thus, WES may be better suited for such cases. Yet, the identification of a single de novo pathogenic mutation in *GALNS* in patient V2.49 can inform genetic counselling in this family. Similarly, the identification of one pathogenic and one potentially pathogenic mutation in *GALNS* in patient V2.15 permits an evidence-based investigation of MPS IVA by assaying leucocyte *GALNS* activity.¹⁰⁴ Therefore, although sequencing did not provide a definitive molecular diagnosis in these cases, the results of this assay can contribute to orienting clinical investigations/counselling.

Among the 21 cases of isolated STHAG, we identified the causative mutation in 15 (71%). This diagnostic rate is comparable to that reported in the literature. By screening *WNT10A*, *MSX1*, *PAX9*, *IRF6* and *AXIN2* by Sanger sequencing in 34 patients with isolated hypodontia, van den Boogaard *et al*⁷⁷ were able to identify the underlying mutation in 71% of cases. Similar to other reports, *WNT10A* accounted for the majority of mutations in our cohort (52%) and were also associated with ectodermal dysplasia (patient V2.54), with F228I, similar to other reports, being the most commonly mutated allele in our cohort, consistent with its population frequency (2.4% in European-Americans) and the prevalence of STHAG in this population (4.2%).^{77–79 105 106}

In conclusion, we have developed the first targeted NGS gene panel for the diagnosis and discovery of mutations in genetic disorders with orodental involvement. This panel can be reliably

used for the molecular diagnosis in known genes of a variety of genetic disorders and can serve as a primary screening tool before the application of WES/WGS. Additionally, this panel also provides the potential for the discovery of novel genes mutated in orodental disorders.

Author affiliations

- ¹Laboratoire de Génétique Médicale, INSERM U1112, Institut de génétique médicale d'Alsace, FMTS, Université de Strasbourg, Strasbourg, France
- ²Plateforme de Biopuces et Séquençage, Institut de Génétique et de Biologie Moléculaire and Cellulaire-Centre Européen de Recherche en Biologie et en Médecine, CNRS UMR7104, INSERM U964, Université de Strasbourg, Illkirch, France
- ³Centre de Référence des Manifestations Odontologiques des Maladies Rares, Pôle de Médecine et Chirurgie Bucco-dentaires, Hôpitaux Universitaires de Strasbourg (HUS), Strasbourg, France
- ⁴Faculté de Chirurgie Dentaire, Université de Strasbourg, Strasbourg, France
- ⁵Evolution et Développement du Squelette-EDS, UMR7138-SAE, Université Pierre et Marie Curie, Paris, France
- ⁶Institut de Génétique et de Biologie Moléculaire and Cellulaire-Centre Européen de Recherche en Biologie et en Médecine, CNRS UMR7104, INSERM U964 Université de Strasbourg, Illkirch, France
- ⁷Faculté de Médecine, CHU de Nancy, Université de Lorraine, Vandoeuvre-Les-Nancy, France
- ⁸CHU de Nancy, Service d'Odontologie, Nancy, France
- ⁹Département d'Odontologie Pédiatrique, UFR d'Odontologie, Université de Nice Sophia-Antipolis, CHU de Nice, Nice, France
- ¹⁰URB2i—EA 4462, Paris Descartes, Paris, France
- ¹¹Faculté de Chirurgie Dentaire, Département d'Odontologie Pédiatrique, CHU Hotel Dieu, Service d'odontologie conservatrice et pédiatrique, Nantes, France
- ¹²Faculté de Chirurgie Dentaire, Département d'Odontologie Pédiatrique, Université Paris Descartes, Montrouge, France
- ¹³Institute of Human Genetics, Heidelberg University, Heidelberg, Germany
- ¹⁴Service de Consultations et Traitements Dentaires, Hospices Civils de Lyon, Faculté d'Odontologie, Université Claude Bernard Lyon1, Lyon, France
- ¹⁵Service de Génétique Clinique, CHU de Rennes, Rennes, France
- ¹⁶Faculté de Chirurgie Dentaire, CHU de Toulouse, Odontologie Pédiatrique, Université Paul Sabatier, Toulouse, France
- ¹⁷Unité Fonctionnelle d'Odontologie pédiatrique, Service d'odontologie, CHRU de Lille, Lille, France
- ¹⁸Faculté de Chirurgie Dentaire, Service d'Odontologie Conservatrice et Endodontie, CHU Nantes, Université de Nantes, France
- ¹⁹Service de Génétique, CHU Tours, Tours, France
- ²⁰Département d'Odontologie Pédiatrique, Université de Nice Sophia-Antipolis, CHU Nice, Nice, France
- ²¹Service de Génétique Médicale, CHU Nice, Nice, France
- ²²Pôle de Médecine et de Chirurgie Bucco-dentaire, Hôpital Civil, HUS, Strasbourg, France
- ²³Faculty of Dental Medicine, Department of Pediatric Dentistry, University Mohammed V Rabat, Morocco
- ²⁴Pédiatrie 1, Hôpitaux Universitaires de Strasbourg, Strasbourg, France
- ²⁵Département de Pédiatrie, CHU de Lyon, Lyon, France
- ²⁶Service de génétique clinique Guy Fontaine, Centre Hospitalier Régionale Universitaire (CHRU) de Lille, Lille, France
- ²⁷Service de Génétique Médicale, CHU de Strasbourg, Strasbourg, France
- ²⁸Klinik für Zahnerhaltungskunde und Parodontologie, Universitäts Klinikum, Freiburg, Germany
- ²⁹Service de Génétique, Centre Hospitalier de Mulhouse, Mulhouse, France
- ³⁰Laboratoire de Physiopathologie Orale Moléculaire INSERM UMR S1138, Centre de Recherche des Cordeliers, Universités Paris-Diderot et Paris-Descartes, Paris, France
- ³¹Centre de Référence des Malformations Rares de la Face et de la Cavité Buccale MAFACE, Hôpital Rothschild, Pôle d'Odontologie, Paris, France
- ³²Aix-Marseille Université, UMR 7268 ADES/EFS/CNRS, APHM, Hôpital Timone, Service Odontologie, Marseille, France
- ³³Service de Génétique Médicale, Centre de Référence pour les Affections Rares en Génétique Ophtalmologique, HUS, Strasbourg, France

Acknowledgements The authors would like to thank all patients and their families for participating in this study. They also thank Ms Vanessa Stoehr for administrative help, Dr Mathilde Huckert and Jorel Salomon for technical assistance, Drs Elise Schaefer and Sophie Scheidecker for help with obtaining clinical data, and Dr Yaumara Perdomo-Trujillo for the electroretinography examination. They would like to thank Drs Alain Verloes, Jean-Yves Sire, Jean-Pierre Clavert, Béatrice Walter, Patrick Lutz, André Nirrengarten, Isabelle Rian, Thibault Sibert, Laurence Lustig-Grimm, Daniel Jasiak, James Lespinasse, Jacques Hassid, Smail Hadj-Rabia, Gabriel Dominici, Laurent Riguet, Dan Lipsker, Jean Luc Alessandri, Florence Rousselet, Philippe Schoenlaub, François Serres, Etienne Mornet, Thomas Edouard,

Bertrand Baumann, Klaus Dietrich, Yves Bolender, Sandra Mercier, Elisabeth Steichen-Gersdorf, Jean-Gabriel Chillès, Hadrien Bonomi Dunoyer, Philippe Jonveaux, Supawich Morkumed, Patimaporn Pungchanchaikul, Sameh Mabrouk, Noora Zouari, Ilhame Mamouni, Alicia Rossa, Caroline Delfosse, Sylvie Goldmann, Esther Levy, Pierre Wahl, Jean-Claude Schoeffler, François Charton, Pascale Meng, Pierre-Emile Spaeth, Marie-Claude Addor, Rene Serfaty, Anthony Kurtz, Renaud Rinkenbach, Alain Duret, Jehan Figier, Denys Chaigne, Marc Hildwein, Florence Jouanet, Olivier Grison, Marie-Noëlle Duffard, Gautier Jouvenet Gilquin, François Undreiner, Gabrielle Baver, Isabelle Zumsteeg, Cloé Vorpahl, Alain Burtcher, Michel Kretz, Cyrielle Marc, Jean-Pierre Strauss, Elise Pilavay, Pierre Bonnaure, and Celine Benarroch-Leininger, Remy Mathis, William Bacon, Marie-Rose Javier, Bilal Ahmed and Elia Sfeir for referring patients. Sequencing was performed by the IGBMC Microarray and Sequencing platform, a member of the 'France Génomique' consortium (ANR-10-INBS-0009).

Contributors Study concept and design: MKP, HD and AB-Z. Patient inclusion: MS, BL, DD, AD, AL, BG, MM-B, SLC, MM, SJ, FO, VV, J-LD, TD-B, A-SK, UM, BR, J-JM, J-PD, SO, IB-F, MMR, LM, AT, CJ, FG, J-CD, AC, MEA, SL, SS, NG, AD, BD, SF, EG, BF, MDM, YA, CT, FC, AB, MCM, HD and AB-Z. Data generation, analysis and interpretation: MKP, VG, SV, BJ, MD, SLG, VL-H, BG, MP, CS and AB-Z. Drafting of the manuscript: MKP and AB-Z. Critical reading of the manuscript: UM, AB and HD.

Funding This work was supported by grants from the French Ministry of Health (National Program for Clinical Research, PHRC 2008 N°4266 Amelogenesis imperfecta), the University Hospital of Strasbourg (API, 2009–2012, 'Development of the oral cavity: from gene to clinical phenotype in Human') and the EU-funded project (ERDF) A27 'Oro-dental manifestations of rare diseases', supported by the RMT-TMO Offensive Sciences initiative, INTERREG IV Upper Rhine programme. <http://www.genosmile.eu>

Competing interests None declared.

Patient consent Obtained.

Ethics approval French Ministry of Higher Education and Research (MESR).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All PCR primers and protocols are available upon request from the corresponding author.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

- Polder BJ, Van't Hof MA, Van Der Linden FP, Kuijpers-Jagtman AM. A meta-analysis of the prevalence of dental agenesis of permanent teeth. *Community Dent Oral Epidemiol* 2004;32:217–26.
- de La Dure-Molla M, Philippe Fournier B, Berdal A. Isolated dentinogenesis imperfecta and dentin dysplasia: revision of the classification. *Eur J Hum Genet* 2015;23:445–51.
- Brook AH. A unifying aetiological explanation for anomalies of human tooth number and size. *Arch Oral Biol* 1984;29:373–8.
- Brook AH, Smith JM. The aetiology of developmental defects of enamel: a prevalence and family study in East London, UK. *Connect Tissue Res* 1998;39:151–6; discussion 87–94.
- Alaluusua S, Lukinmaa PL. Developmental dental toxicity of dioxin and related compounds—a review. *Int Dent J* 2006;56:323–31.
- Koch G. Prevalence of enamel mineralisation disturbances in an area with 1–1.2 ppm F in drinking water. Review and summary of a report published in Sweden in 1981. *Eur J Paediatr Dent* 2003;4:127–8.
- Bloch-Zupan A, Sedano H, Scully C. *Dento/oro/craniofacial anomalies and genetics*. 1st edn. Boston, MA: Elsevier, 2012.
- Cobourne MT, Sharpe PT. Diseases of the tooth: the genetic and molecular basis of inherited anomalies affecting the dentition. *Wiley Interdiscip Rev Dev Biol* 2013;2:183–212.
- Crawford PJ, Aldred M, Bloch-Zupan A. Amelogenesis imperfecta. *Orphanet J Rare Dis* 2007;2:17.
- Hennekam RCM, Allanson JE, Krantz ID, Gorlin RJ. *Gorlin's syndromes of the head and neck*. 5th edn. Oxford; New York: Oxford University Press, 2010.
- Bergendal B, Klar J, Stecksén-Blicks C, Norderyd J, Dahl N. Isolated oligodontia associated with mutations in EDARADD, AXIN2, MSX1, and PAX9 genes. *Am J Med Genet A* 2011;155A:1616–22.
- Lammi L, Arte S, Somer M, Järvinen H, Lahermo P, Thesleff I, Pirinen S, Nieminen P. Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. *Am J Hum Genet* 2004;74:1043–50.
- Noor A, Windpassinger C, Vitcu I, Orlic M, Arshad Rafiq M, Khalid M, Nasir Malik M, Ayub M, Alman B, Vincent JB. Oligodontia is caused by mutation in LTBP3, the gene encoding latent TGF-beta binding protein 3. *Am J Hum Genet* 2009;84:519–23.
- Stockton DW, Das P, Goldenberg M, D'souza RN, Patel PI. Mutation of PAX9 is associated with oligodontia. *Nat Genet* 2000;24:18–19.
- Tao R, Jin B, Guo SZ, Qing W, Feng GY, Brooks DG, Liu L, Xu J, Li T, Yan Y, He L. A novel missense mutation of the EDA gene in a Mongolian family with congenital hypodontia. *J Hum Genet* 2006;51:498–502.
- Vastardis H, Karimlux N, Guthua SW, Seidman JG, Seidman CE. A human MSX1 homeodomain missense mutation causes selective tooth agenesis. *Nat Genet* 1996;13:417–21.
- Monreal AW, Zonana J, Ferguson B. Identification of a new splice form of the EDA1 gene permits detection of nearly all X-linked hypohidrotic ectodermal dysplasia mutations. *Am J Hum Genet* 1998;63:380–9.
- Cluzeau C, Hadj-Rabia S, Jambou M, Mansour S, Guigue P, Masmoudi S, Bal E, Chassaing N, Vincent M-C, Viot G, Clauss F, Manière M-C, Toupenay S, Le Merrer M, Lyonnet S, Cormier-Daire V, Amiel J, Faivre L, De Prost Y, Munnich A, Bonnefont J-P, Bodemer C, Smahi A. Only four genes (EDA1, EDAR, EDARADD, and WNT10A) account for 90% of hypohidrotic/anhidrotic ectodermal dysplasia cases. *Hum Mutat* 2011;32:70–2.
- Jumlongras D, Bei M, Stimson JM, Wang W-F, Depalma SR, Seidman CE, Felbor U, Maas R, Seidman JG, Olsen BR. A nonsense mutation in MSX1 causes Witkop syndrome. *Am J Hum Genet* 2001;69:67–74.
- Monreal AW, Ferguson BM, Headon DJ, Street SL, Overbeek PA, Zonana J. Mutations in the human homologue of mouse dl cause autosomal recessive and dominant hypohidrotic ectodermal dysplasia. *Nat Genet* 1999;22:366–9.
- Rajpar MH, Harley K, Laing C, Davies RM, Dixon MJ. Mutation of the gene encoding the enamel-specific protein, enamelin, causes autosomal-dominant amelogenesis imperfecta. *Hum Mol Genet* 2001;10:1673–7.
- El-Sayed W, Parry DA, Shore RC, Ahmed M, Jafri H, Rashid Y, Al-Bahlani S, Al Harasi S, Kirkham J, Inglehearn CF, Mighell AJ. Mutations in the beta propeller WDR72 cause autosomal-recessive hypomaturation amelogenesis imperfecta. *Am J Hum Genet* 2009;85:699–705.
- Hart PS, Hart TC, Michalec MD, Ryu OH, Simmons D, Hong S, Wright JT. Mutation in kallikrein 4 causes autosomal recessive hypomaturation amelogenesis imperfecta. *J Med Genet* 2004;41:545–9.
- Lagerström M, Dahl N, Nakahori Y, Nakagome Y, Bäckman B, Landegren U, Pettersson U. A deletion in the amelogenin gene (AMG) causes X-linked amelogenesis imperfecta (AIH1). *Genomics* 1991;10:971–5.
- Kim JW, Simmer JP, Hart TC, Hart PS, Ramaswami MD, Bartlett JD, Hu JC. MMP-20 mutation in autosomal recessive pigmented hypomaturation amelogenesis imperfecta. *J Med Genet* 2005;42:271–5.
- Kim JW, Lee SK, Lee ZH, Park J-C, Lee K-E, Lee M-H, Park J-T, Seo B-M, Hu JC-C, Simmer JP. FAM83H mutations in families with autosomal-dominant hypocalcified amelogenesis imperfecta. *Am J Hum Genet* 2008;82:489–94.
- Poulter JA, Murillo G, Brookes SJ, Smith CEL, Parry DA, Silva S, Kirkham J, Inglehearn CF, Mighell AJ. Deletion of ameloblastin exon 6 is associated with amelogenesis imperfecta. *Hum Mol Genet* 2014;23:5317–24.
- Poulter JA, Brookes SJ, Shore RC, Smith CEL, Abi Farraj L, Kirkham J, Inglehearn CF, Mighell AJ. A missense mutation in ITGB6 causes pitted hypomineralized amelogenesis imperfecta. *Hum Mol Genet* 2014;23:2189–97.
- Parry DA, Poulter JA, Logan CV, Brookes SJ, Jafri H, Ferguson CH, Anwari BM, Rashid Y, Zhao H, Johnson CA, Inglehearn CF, Mighell AJ. Identification of mutations in SLC24A4, encoding a potassium-dependent sodium/calcium exchanger, as a cause of amelogenesis imperfecta. *Am J Hum Genet* 2013;92:307–12.
- Parry DA, Brookes SJ, Logan CV, Poulter JA, El-Sayed W, Al-Bahlani S, Al Harasi S, Sayed J, Raif EM, Shore RC, Dashash M, Barron M, Morgan JE, Carr IM, Taylor GR, Johnson CA, Aldred MJ, Dixon MJ, Wright JT, Kirkham J, Inglehearn CF, Mighell AJ. Mutations in C4orf26, encoding a peptide with in vitro hydroxyapatite crystal nucleation and growth activity, cause amelogenesis imperfecta. *Am J Hum Genet* 2012;91:565–71.
- Huckert M, Stoetzel C, Morkmued S, Laugel-Hausalter V, Geoffroy V, Muller J, Clauss F, Prasad MK, Obyr F, Raymond JL, Switala M, Alembik Y, Soskin S, Mathieu E, Hemmerle J, Weickert J-L, Dabovic BB, Rifkin DB, Dheedene A, Boudin E, Caluseriu O, Cholette M-C, McLeod R, Antequera R, Gelle M-P, Coeuriot J-L, Jacquelin L-F, Bailleul-Forestier I, Manière M-C, Van Hul W, Bertola D, Dolle P, Verloes A, Mortier G, Dollfus H, Bloch-Zupan A. Mutations in the Latent TGF-beta Binding Protein 3 (LTBP3) gene cause brachyolmia with amelogenesis imperfecta. *Hum Mol Genet* 2015;24:3038–49.
- O'Sullivan J, Bitu CC, Daly SB, Urquhart JE, Barron MJ, Bhaskar SS, Martelli-Júnior H, dos Santos Neto PE, Mansilla MA, Murray JC, Coletta RD, Black GCM, Dixon MJ. Whole-Exome sequencing identifies FAM20A mutations as a cause of amelogenesis imperfecta and gingival hyperplasia syndrome. *Am J Hum Genet* 2011;88:616–20.
- Parry DA, Mighell AJ, El-Sayed W, Shore RC, Jalili IK, Dollfus H, Bloch-Zupan A, Carlos R, Carr IM, Downey LM, Blain KM, Mansfield DC, Shahrabi M, Heidari M,

- Aref P, Abbasi M, Michaelides M, Moore AT, Kirkham J, Inglehearn CF. Mutations in CNM4 cause Jalili syndrome, consisting of autosomal-recessive cone-rod dystrophy and amelogenesis imperfecta. *Am J Hum Genet* 2009;84:266–73.
- 34 Schossig A, Wolf NI, Fischer C, Fischer M, Stocker G, Pabinger S, Dander A, Steiner B, Tönz O, Kotzot D, Haberlandt E, Amberger A, Burwinkel B, Wimmer K, Fauth C, Grond-Ginsbach C, Koch MJ, Deichmann A, von Kalle C, Bartram CR, Kohlschütter A, Trajanoski Z, Zschocke J. Mutations in ROGDI Cause Kohlschütter-Tönz Syndrome. *Am J Hum Genet* 2012;90:701–7.
- 35 Acevedo AC, Poulter JA, Alves PG, De Lima C, Castro L, Yamaguti P, Paula LM, Parry DA, Logan CV, Smith CEL, Johnson CA, Inglehearn CF, Mighell AJ. Variability of systemic and oro-dental phenotype in two families with non-lethal Raine syndrome with FAM20C mutations. *BMC Med Genet* 2015;16:8.
- 36 Wang S, Choi M, Richardson AS, Reid BM, Seymen F, Yildirim M, Tuna E, Gencay K, Simmer JP, Hu JC. STIM1 and SLC24A4 Are Critical for Enamel Maturation. *J Dent Res* 2014;93(7 Suppl):945–1005.
- 37 Mcgrath JA, Gatalica B, Li K, Dunnill MG, McMillan JR, Christiano AM, Eady RA, Uitto J. Compound heterozygosity for a dominant glycine substitution and a recessive internal duplication mutation in the type XVII collagen gene results in junctional epidermolysis bullosa and abnormal dentition. *Am J Pathol* 1996;148:1787–96.
- 38 Poulter JA, El-Sayed W, Shore RC, Kirkham J, Inglehearn CF, Mighell AJ. Whole-exome sequencing, without prior linkage, identifies a mutation in LAMB3 as a cause of dominant hypoplastic amelogenesis imperfecta. *Eur J Hum Genet* 2014;22:132–5.
- 39 Yuen WY, Pasmooij AM, Stellingma C, Jonkman MF. Enamel defects in carriers of a novel LAMA3 mutation underlying epidermolysis bullosa. *Acta Derm Venereol* 2012;92:695–6.
- 40 Dong J, Amor D, Aldred MJ, Gu T, Escamilla M, Macdougall M. DLX3 mutation associated with autosomal dominant amelogenesis imperfecta with taurodontism. *Am J Med Genet A* 2005;133A:138–41.
- 41 de La Dure-Molla M, Quentric M, Yamaguti PM, Acevedo A-C, Mighell AJ, Viikula M, Huckert M, Berdal A, Bloch-Zupan A. Pathognomonic oral profile of Enamel Renal Syndrome (ERS) caused by recessive FAM20A mutations. *Orphanet J Rare Dis* 2014;9:84.
- 42 Chae JH, Vasta V, Cho A, Lim BC, Zhang Q, Eun SH, Hahn SH. Utility of next generation sequencing in genetic diagnosis of early onset neuromuscular disorders. *J Med Genet* 2015;52:208–16.
- 43 DaRe JT, Vasta V, Penn J, Tran NT, Hahn SH. Targeted exome sequencing for mitochondrial disorders reveals high genetic heterogeneity. *BMC Med Genet* 2013;14:118.
- 44 Redin C, Gérard B, Lauer J, Herenger Y, Muller J, Quartier A, Masurel-Paulet A, Willems M, Lesca G, El-Chehadeh S, Le Gras S, Vicaire S, Philipps M, Dumas M, Geoffroy V, Feger C, Haumesser N, Alembik Y, Barth M, Bonneau D, Colin E, Dollfus H, Doray B, Delrue M-A, Drouin-Garraud V, Flori E, Fradin M, Francannet C, Goldenberg A, Lumbroso S, Mathieu-Dramard M, Martin-Coignard D, Lacombe D, Morin G, Polge A, Sukno S, Thauvin-Robinet C, Thevenon J, Docco-Fenzy M, Genevieve D, Sarda P, Ederly P, Isidor B, Jost B, Olivier-Faivre L, Mandel J-L, Piton A. Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. *J Med Genet* 2014;51:724–36.
- 45 Redin C, Le Gras S, Mhamdi O, Geoffroy V, Stoetzel C, Vincent M-C, Chiurazzi P, Lacombe D, Ouertani I, Petit F, Till M, Verloes A, Jost B, Chaabouni HB, Dollfus H, Mandel J-L, Muller J. Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alström syndromes. *J Med Genet* 2012;49:502–12.
- 46 Shearer AE, DeLuca AP, Hildebrand MS, Taylor KR, Gurrrola J, Scherer S, Scheetz TE, Smith RJH. Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. *Proc Natl Acad Sci USA* 2010;107:21104–9.
- 47 Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet* 2013;14:295–300.
- 48 Eppig JT, Blake JA, Bult CJ, Kadin JA, Richardson JE, Mouse Genome Database G. The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease. *Nucleic Acids Res* 2015;43(Database issue):D726–36.
- 49 Bradford Y, Conlin T, Dunn N, Fashena D, Frazer K, Howe DG, Knight J, Mani P, Martin R, Moxon SAT, Paddock H, Pich C, Ramachandran S, Ruef BJ, Ruzicka L, Bauer Schaper H, Schaper K, Shao X, Singer A, Sprague J, Sprunger B, Van Slyke C, Westerfield M. ZFIN: enhancements and updates to the Zebrafish Model Organism Database. *Nucleic Acids Res* 2011;39(Database issue):D822–9.
- 50 Laugel-Haushalter V, Paschaki M, Thibault-Carpentier C, Dembelé D, Dollé P, Bloch-Zupan A. Molars and incisors: show your microarray IDs. *BMC Res Notes* 2013;6:113.
- 51 Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;26:589–95.
- 52 DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, Del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernysky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491–8.
- 53 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
- 54 Van Der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013;11:11 10 1–11 10 33.
- 55 Cingolani P, Platts A, Wang Le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 2012;6:80–92.
- 56 Geoffroy V, Pizot C, Redin C, Piton A, Vasli N, Stoetzel C, Blavier A, Laporte J, Muller J. VaRank: a simple and powerful tool for ranking genetic variants. *PeerJ* 2015;3:e796.
- 57 Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA, 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012;491:56–65.
- 58 Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA. <http://evs.gs.washington.edu/EVS/> (accessed Feb 2015).
- 59 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
- 60 Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073–81.
- 61 Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* 2014;11:361–2.
- 62 Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 2003;31:3568–71.
- 63 Desmet FO, Hamroun D, Lalonde M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 2009;37:e67.
- 64 Fairbrother WG, Yeh RF, Sharp PA, Burge CB. Predictive identification of exonic splicing enhancers in human genes. *Science* 2002;297:1007–13.
- 65 Perteau M, Lin X, Salzberg SL. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res* 2001;29:1185–90.
- 66 Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol* 1997;4:311–23.
- 67 Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 2004;11:377–94.
- 68 Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, Lyon E, Ward BE. ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. *Genet Med* 2008;10:294–300.
- 69 Kravtsov Y, Nekhaeva E, Chang C, Ebralidse K, Khrapko K. Analysis of somatic mutations via long-distance single molecule PCR. *DNA Amplif* 2004:97–110.
- 70 Ruano G, Kidd KK, Stephens JC. Haplotype of multiple polymorphisms resolved by enzymatic amplification of single DNA molecules. *Proc Natl Acad Sci USA* 1990;87:6296–300.
- 71 Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001;29:e45.
- 72 Seymen F, Lee KE, Tran Le CG, Yildirim M, Gencay K, Lee ZH, Kim J-W. Exonal deletion of SLC24A4 causes hypomaturational amelogenesis imperfecta. *J Dent Res* 2014;93:366–70.
- 73 Darling TN, Mcgrath JA, Yee C, Gatalica B, Hametner R, Bauer JW, Pohla-Gubo G, Christiano AM, Uitto J, Hintner H, Yancey KB. Premature termination codons are present on both alleles of the bullous pemphigoid antigen 2/type XVII collagen gene in five Austrian families with generalized atrophic benign epidermolysis bullosa. *J Invest Dermatol* 1997;108:463–8.
- 74 Mcgrath JA, Pulkkinen L, Christiano AM, Leigh IM, Eady RA, Uitto J. Altered laminin 5 expression due to mutations in the gene encoding the beta3 chain (LAMB3) in generalized atrophic benign epidermolysis bullosa. *J Invest Dermatol* 1995;104:467–74.
- 75 Wright JT, Torain M, Long K, Seow K, Crawford P, Aldred MJ, Hart PS, Hart TC. Amelogenesis imperfecta: genotype-phenotype studies in 71 families. *Cells Tissues Organs* 2011;194:279–83.
- 76 Lee SK, Hu JC, Bartlett JD, Lee K-E, Lin BP-J, Simmer JP, Kim J-W. Mutational spectrum of FAM83H: the C-terminal portion is required for tooth enamel calcification. *Hum Mutat* 2008;29:E95–9.
- 77 van den Boogaard MJ, Créton M, Bronkhorst Y, Van Der Hout A, Hennekam E, Lindhout D, Cune M, Ploos Van Amstel HK. Mutations in WNT10A are present in more than half of isolated hypodontia cases. *J Med Genet* 2012;49:327–31.
- 78 Bohring A, Stamm T, Spaich C, Haase C, Spree K, Hehr U, Hoffmann M, Ledig S, Sel S, Wieacker P, Röpke A. WNT10A mutations are a frequent cause of a broad spectrum of ectodermal dysplasias with sex-biased manifestation pattern in heterozygotes. *Am J Hum Genet* 2009;85:97–105.

- 79 Plaisancié J, Bailleul-Forestier I, Gaston V, Vaysse F, Lacombe D, Holder-Espinasse M, Abramowicz M, Coubes C, Plessis G, Faivre L, Demeer B, Vincent-Delorme C, Dollfus H, Sigaudy S, Guillén-Navarro E, Verloes A, Jonveaux P, Martin-Coignard D, Colin E, Bieth E, Calvas P, Chassaing N. Mutations in WNT10A are frequently involved in oligodontia associated with minor signs of ectodermal dysplasia. *Am J Med Genet A* 2013;161:671–8.
- 80 Xiao S, Yu C, Chou X, Yuan W, Wang Y, Bu L, Fu G, Qian M, Yang J, Shi Y, Hu L, Han B, Wang Z, Huang W, Liu J, Chen Z, Zhao G, Kong X. Dentinogenesis imperfecta 1 with or without progressive hearing loss is associated with distinct mutations in DSPP. *Nat Genet* 2001;27:201–4.
- 81 Wang Z, Zhang W, Wang Y, Meng Y, Su L, Shi H, Huang S. Mucopolysaccharidosis IVA mutations in Chinese patients: 16 novel mutations. *J Hum Genet* 2010;55:534–40.
- 82 Cho SH, Seymen F, Lee KE, Lee S-K, Kweon Y-S, Kim KJ, Jung S-E, Song SJ, Yildirim M, Bayram M, Tuna EB, Gencyay K, Kim J-W. Novel FAM20A mutations in hypoplastic amelogenesis imperfecta. *Hum Mutat* 2012;33:91–4.
- 83 Kivirikko S, McGrath JA, Pulkkinen L, Uitto J, Christiano AM. Mutational hotspots in the LAMB3 gene in the lethal (Herlitz) type of junctional epidermolysis bullosa. *Hum Mol Genet* 1996;5:231–7.
- 84 Ogawa T, Tomatsu S, Fukuda S, Yamagishi A, Rezvi GMM, Sukegawa K, Kondo N, Suzuki Y, Shimozawa N, Orii T. Mucopolysaccharidosis IVA: screening and identification of mutations of the N-acetylgalactosamine-6-sulfate sulfatase gene. *Hum Mol Genet* 1995;4:341–9.
- 85 Pasmooij AM, Pas HH, Jansen GH, Lemmink HH, Jonkman MF. Localized and generalized forms of blistering in junctional epidermolysis bullosa due to COL17A1 mutations in the Netherlands. *Br J Dermatol* 2007;156:861–70.
- 86 Murrell DF, Pasmooij AM, Pas HH, Marr P, Klingberg S, Pfendner E, Uitto J, Sadowski S, Collins F, Widmer R, Jonkman MF. Retrospective diagnosis of fatal BP180-deficient non-Herlitz junctional epidermolysis bullosa suggested by immunofluorescence (IF) antigen-mapping of parental carriers bearing enamel defects. *J Invest Dermatol* 2007;127:1772–5.
- 87 Floeth M, Bruckner-Tuderman L. Digenic junctional epidermolysis bullosa: mutations in COL17A1 and LAMB3 genes. *Am J Hum Genet* 1999;65:1530–7.
- 88 Beyeler M, Schild C, Lutz R, Chiquet M, Trueb B. Identification of a fibronectin interaction site in the extracellular matrix protein ameloblastin. *Exp Cell Res* 2010;316:1202–12.
- 89 Zhang X, Diekwisch TG, Luan X. Structure and function of ameloblastin as an extracellular matrix protein: adhesion, calcium binding, and CD63 interaction in human and mouse. *Eur J Oral Sci* 2011;119(Suppl 1):270–9.
- 90 Polok B, Escher P, Ambresin A, Chouery E, Bolay S, Meunier I, Nan F, Hamel C, Munier FL, Thilo B, Mégarbané A, Schorderet DF. Mutations in CNNM4 cause recessive cone-rod dystrophy with amelogenesis imperfecta. *Am J Hum Genet* 2009;84:259–65.
- 91 Mendelsohn NJ, Wood T, Olson RA, Temme R, Hale S, Zhang H, Read L, White KK. Spondyloepiphyseal dysplasias and bilateral legg-calvé-perthes disease: diagnostic considerations for mucopolysaccharidoses. *JIMD Rep* 2013;11:125–32.
- 92 Gardner DG. The dental manifestations of the Morquio syndrome (mucopolysaccharidosis type IV). A diagnostic aid. *Am J Dis Child* 1975;129:1445–8.
- 93 Levin LS, Jorgenson RJ, Salinas CF. Oral findings in the Morquio syndrome (mucopolysaccharidosis IV). *Oral Surg Oral Med Oral Pathol* 1975;39:390–5.
- 94 Rølling I, Clausen N, Nyvad B, Sindet-Pedersen S. Dental findings in three siblings with Morquio's syndrome. *Int J Paediatr Dent* 1999;9:219–24.
- 95 Liu F, Chu EY, Watt B, Zhang Y, Gallant NM, Andl T, Yang SH, Lu M-M, Piccolo S, Schmidt-Ullrich R, Taketo MM, Morrissy EE, Atit R, Dlugosz AA, Millar SE. Wnt/ beta-catenin signaling directs multiple stages of tooth morphogenesis. *Dev Biol* 2008;313:210–24.
- 96 de Ligt J, Willemssen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-Van Silfhout AT, Koolen DA, De Vries P, Gilissen C, Del Rosario M, Hoischen A, Scheffer H, De Vries BBA, Brunner HG, Veltman JA, Vissers LELM. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 2012;367:1921–9.
- 97 Tucci V, Kleefstra T, Hardy A, Heise I, Maggi S, Willemssen MH, Hilton H, Esapa C, Simon M, Buenavista M-T, Mcguffin LJ, Vizer L, Dodero L, Tsaftaris S, Romero R, Nillesen WN, Vissers LELM, Kempers MJ, Vulto-Van Silfhout AT, Iqbal Z, Orlando M, Maccione A, Lassi G, Farisello P, Contestabile A, Tinarelli F, Nieuws T, Raimondi A, Greco B, Cantatore D, Gasparini L, Berdondini L, Bifone A, Gozzi A, Wells S, Nolan PM. Dominant beta-catenin mutations cause intellectual disability with recognizable syndromic features. *J Clin Invest* 2014;124:1468–82.
- 98 McKnight DA, Suzanne Hart P, Hart TC, Hartsfield JK, Wilson A, Wright JT, Fisher LW. A comprehensive analysis of normal variation and disease-causing mutations in the human DSPP gene. *Hum Mutat* 2008;29:1392–404.
- 99 Gregory-Evans CY, Moosajee M, Hodges MD, Mackay DS, Game L, Vargesson N, Bloch-Zupan A, Ruschendorf F, Santos-Pinto L, Wackens G, Gregory-Evans K. SNP genome scanning localizes oto-dental syndrome to chromosome 11q13 and microdeletions at this locus implicate FGF3 in dental and inner-ear disease and FADD in ocular coloboma. *Hum Mol Genet* 2007;16:2482–93.
- 100 Grozeva D, Carss K, Spasic-Boskovic O, Parker MJ, Archer H, Firth HV, Park S-M, Canham N, Holder SE, Wilson M, Hackett A, Field M, Floyd JAB, Hurles M, Raymond FL. De novo loss-of-function mutations in SETD5, encoding a methyltransferase in a 3p25 microdeletion syndrome critical region, cause intellectual disability. *Am J Hum Genet* 2014;94:618–24.
- 101 Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, Mcmillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* 2010;42:790–3.
- 102 Chan HC, Estrella NM, Milkovich RN, Kim JW, Simmer JP, Hu JC. Target gene analyses of 39 amelogenesis imperfecta kindreds. *Eur J Oral Sci* 2011;119(Suppl 1):311–23.
- 103 Gasse B, Karayigit E, Mathieu E, Jung S, Garret A, Huckert M, Morkmued S, Schneider C, Vidal L, Hemmerle J, Sire J-Y, Bloch-Zupan A. Homozygous and compound heterozygous MMP20 mutations in amelogenesis imperfecta. *J Dent Res* 2013;92:598–603.
- 104 Wood TC, Harvey K, Beck M, Burin MG, Chien Y-H, Church HJ, D'Almeida V, van Diggelen OP, Fietz M, Giugliani R, Harmatz P, Hawley SM, Hwu W-L, Ketteridge D, Lukacs Z, Miller N, Pasquali M, Schenone A, Thompson JN, Tylee K, Yu C, Hendriksz CJ. Diagnosing mucopolysaccharidosis IVA. *J Inherit Metab Dis* 2013;36:293–307.
- 105 Arzoo PS, Klar J, Bergendal B, Norderyd J, Dahl N. WNT10A mutations account for (1/4) of population-based isolated oligodontia and show phenotypic correlations. *Am J Med Genet A* 2014;164A:353–9.
- 106 Mostowska A, Biedziak B, Zadurska M, Dunin-Wilczynska I, Lianeri M, Jagodzinski PP. Nucleotide variants of genes encoding components of the Wnt signalling pathway and the risk of non-syndromic tooth agenesis. *Clin Genet* 2013;84:429–40.