PAX9 mutations and genetic synergism in familial tooth agenesis

Kuan-Yu Chu[1,2], Yin-Lin Wang[1,2], Jung-Tsu Chen[1], Chia-Hui Lin[1], Chung-Chen Jane Yao[1], Yi-Jane Chen[1], Huan-Wen Chen[2], James P. Simmer[3], Jan C.-C. Hu[3], Shih-Kai Wang[1,2*]

[1]Department of Dentistry, National Taiwan University School of Dentistry, No.1, Taipei City, Taiwan.

[2]Department of Pediatric Dentistry, National Taiwan University Children's Hospital, Taipei City, Taiwan.

[3]Department of Biologic and Materials Sciences, University of Michigan School of Dentistry, Ann Arbor, Michigan, USA.

1) Short Title: *PAX9* Mutations and FTA

2) Abstract word count: 200

3) Total word count (Abstract to Acknowledgments): 3681

4) Number of Figures: 4

5) Number of tables: 1

6) Number of cited references: 36

7) Key Words: oligodontia, hypodontia, WNT10A, Wnt signaling, genotype–phenotype correlation, transcription factor

***Corresponding author:**

Shih-Kai Wang DDS, PhD

Department of Dentistry, National Taiwan University School of Dentistry, No.1, Changde St., Jhongjheng District, Taipei City 100229, Taiwan.

Email: shihkaiw@ntu.edu.tw

Graphical abstract

Familial tooth agenesis (FTA) is one of the most common craniofacial anomalies in humans. Loss-of-function mutations in *PAX9* and *WNT10A* have been known to cause FTA with various expressivity. In this study, we identified 5 FTA kindreds with novel *PAX9* disease-causing mutations. Concomitant *PAX9* and *WNT10A* pathogenic variants found in two probands with severe phenotypes suggested an effect of mutational synergism.

ABSTRACT

Familial tooth agenesis (FTA) is one of the most common craniofacial anomalies in humans. Loss-of-function mutations in *PAX9* and *WNT10A* have been known to cause FTA with various expressivity. In this study, we identified 5 FTA kindreds with novel *PAX9* disease-causing mutations, p.(Glu7Lys), p.(Val83Leu), p.(Pro118Ser), p.(Ser197Argfs*23), and c.771+4A>G. Concomitant *PAX9* and *WNT10A* pathogenic variants found in two probands with severe phenotypes suggested an effect of mutational synergism. All overexpressed PAX9s showed proper nuclear localization,

excepting the p.(Pro118Ser) mutant. Various missense mutations caused differential loss of *PAX9* transcriptional ability. *PAX9* overexpression in dental pulp cells upregulated *LEF1* and *AXIN2* expression, indicating a positive regulatory role for PAX9 in canonical Wnt signaling. Analyzing 176 cases with 63 different mutations, we observed a distinct pattern of tooth agenesis for *PAX9*-associated FTA: Maxillary teeth are in general more frequently affected than mandibular ones. Along with all second molars, maxillary bicuspids and first molars are mostly involved, while maxillary lateral incisors and mandibular bicuspids are relatively less affected. Genotypically, missense mutations are associated with fewer missing teeth than frameshift and nonsense variants. This study significantly expands the phenotypic and genotypic spectrums of *PAX9*-associated disorders and reveals a molecular mechanism of genetic synergism underlying FTA variable expressivity.

INTRODUCTION

Familial tooth agenesis (FTA) is a group of genetic disorders characterized by developmental absence of teeth, with minor or no nondental manifestations.[1] FTA is one of the most common craniofacial anomalies in humans and imposes significant esthetic and functional burdens in severe cases.[2] Clinically, hypodontia and oligodontia are used to describe the condition of 5-or-fewer and 6-or-more missing teeth, respectively. Mutations in several genes have been proven to cause non-syndromic FTA, including *MSX1*,[3] *PAX9*,[4] *AXIN2*,[5] *EDA*,[6] *WNT10A*,[7] and *LRP6*.[8] Among these candidate genes, *PAX9* (OMIM*167416) is the one that has been shown to be associated with only dental phenotypes, while mutations in the others can cause syndromes with tooth agenesis, such as ectodermal dysplasia.[9, 10] It has been reported that FTAs caused by different mutated genes appear to have distinct patterns of tooth agenesis.[9, 11] In *PAX9*-associated FTA, affected individuals usually exhibit severe

oligodontia and have agenesis of all molars and mandibular central incisors.[12] However, this pattern of missing teeth requires further scrutinization and validation, and potential genotype–phenotype correlations of *PAX9*-assciated FTA remain to be elucidated.

Variable expressivity is frequently observed in FTA.[1] Even family members carrying the same disease-causing mutation within a family may vary in the severity of their tooth agenesis. It has been considered that the variation in expression of tooth agenesis can be attributed to genetic modifiers, epigenetic regulation, and environmental factors.[1] Potential digenic and multigenic inheritances have also been proposed due to the high genetic heterogeneity of FTA.[13, 14] Genetic synergism refers to a super-additive effect between two or more mutant phenotypes.[15] It occurs when combination of two or more mutations result in a phenotype exceeding the simple summed effect of the individual mutations, which can contribute to variable expressivity of an inherited disorder with genetic heterogeneity, such as FTA. Oligodontia patients who carried both *WNT10A*- and *EDA*-related mutant alleles exhibit a more severe phenotype when compared with those carrying either mutation alone, suggesting a synergistic effect on expressivity and an interaction between Wnt and EDA signaling pathways during tooth development.[16] However, this genetic synergy in FTA has not been fully explored. In this study, we characterized 5 FTA kindreds and, in each of them, identified an unreported *PAX9* disease-causing mutation. The two probands with severe oligodontia carried not only *PAX9* mutations but *WNT10A* pathogenic variants, demonstrating genetic synergy in FTA. Conducting a through literature review, we further characterized the specific pattern of tooth agenesis in *PAX9*-associated FTA and revealed potential genotype–phenotype correlations.

MATERIALS AND METHODS

Recruitment of FTA families and mutational analyses

All the recruitment procedures were specified in our human study protocols and followed the Helsinki Declaration. The protocols and consent forms were reviewed and approved by the IRB Committee at the National Taiwan University Hospital. All study participants signed written consent forms after comprehensive explanation and discussion of the research content. Phenotypic examinations were conducted for disease characterization and pedigree construction. A 2-ml sample of nonstimulated saliva was collected from each subject to obtain genomic DNA for mutational analyses.

To search for FTA-causing mutations, whole exome sequencing and analysis were conducted for each proband as previously described.[14] Target amplification and Sanger sequencing were further performed for mutation validation and segregation analyses for all participants using the corresponding primer sets for *PAX9* and *WNT10A*.

Minigene splicing assay

Wild-type and mutant *PAX9* gene fragments of 832 bp, encompassing exon 4 and parts of introns 3 and 4, were amplified from the genomic DNA of the Family 1 proband and cloned into the pSPL3 vector (a gift from Dr. Tompson) using *Xho*I and *Bam*HI restriction enzyme sites.[17] These two minigene constructs were introduced into COS7 cells for splicing analysis. Twenty-four hours after transfection, RNA was extracted from harvested cells, reverse-transcribed, and amplified with V1-F and V2-R primers. The PCR products were analyzed by agarose gel electrophoresis and DNA sequencing.

Preparation of expression constructs

The human *PAX9* coding region (CCDS9662.1) was synthesized and subcloned into the pcDNA3.1/myc-His(-)A expression vector (ThermoFisher), using *Nhe*I(5′) and *Hind*III(3′) restriction enzyme sites to express N-terminal Myc-tagged human PAX9. Site-directed mutagenesis of the *PAX9* wild-type sequence was used to generate three constructs overexpressing mutant PAX9 proteins p.Glu7Lys, p.Val83Leu, and p.Pro118Ser. For the luciferase assay, two reporter gene constructs were made using pGL4.10[*luc2*] vector (Promega). One was driven by a minimal promoter following six modified human CD19 (CD19-A-ins) BSAP-binding sites,[18] and the other by the human *BMP4* promoter (-2742~+252). The pGL4.74[*hRluc/TK*] vector (Promega) was employed as an internal expression control.

Immunofluorescent staining and dual-luciferase reporter assay

The immunofluorescence and luciferase reporter assays were conducted on COS7 cells as previously described.[19] For luciferase assays, equal amounts of mutant (or empty vector) and wild-type *PAX9* plasmids were co-transfected into the cells to mimic the heterozygous condition in FTA. For immunofluorescence, the mouse monoclonal anti-c-Myc antibody (A00704S; GenScript) was employed as the primary antibody, and goat anti-mouse IgG2a antibody conjugated with Alexa Fluor™ 594 (A-21135; ThermoFisher) as the secondary antibody.

Culture of dental pulp cells, transfection, and RT-PCR

Human dental pulp cells were harvested from a natal tooth of a 5-day-old baby and cultured in DMEM medium with 10% fetal bovine serum at 37°C and 5% CO_2. Cells from passage 4 to 6 were used for the experiments. For overexpression, ~5 × 10^5 cells were seeded in each well of a 6-well plate and transfected with wild-type and p.Pro118Ser *PAX9* expression constructs along with the empty vector control. After 48 hours, cells were harvested and lysed, and RNA extracted. Expression of *PAX9*, *LEF1*, *AXIN2*, *BMP4*, *MSX1*, and *GAPDH* was analyzed by RT-PCR with specific primers (Supplementary Methods) and quantified using ImageJ software.

Literature review

A systematic PubMed/MEDLINE and Google search from 1998 to 2022 identified 57 English articles reporting tooth agenesis phenotypes and *PAX9* genotypes. Only 37 publications were included for further statistical analyses. Criteria for manuscript exclusion, and statistical methods are detailed in the Supplementary Methods.

Statistical analyses for genotype–phenotype correlation

For genotype–phenotype analyses, 37 articles were retained and scrutinized as they provided a phenotypic description of tooth agenesis for at least one individual carrying *PAX9* sequence variants. For each individual carrying a distinct *PAX9* mutation (M, N, F, or U), the numbers of missing teeth in total, in segments (anterior and posterior), and in each tooth type, excluding third molars, were calculated, with numbers from the left and right sides being pooled together. One-way ANOVA (analysis of variance) followed by post-hoc tests was used to evaluate the mean numbers of missing

teeth among individuals with four *PAX9* mutation types. In given analyses for specific mutation groups, percentages of missing teeth for each tooth type (U1–U7 for maxillary and L1–L7 for mandibular teeth) were calculated as the quotient of missing tooth number and total tooth number from all individuals within the group. Pearson's chi-squared tests were conducted to assess the association between absence of specific tooth types and different mutation groups (M, N, F, and U; NMD and non-NMD; inside and outside PD; PAI and RED subdomains). For evaluating the differential involvement of maxillary and mandibular teeth, paired sample *t*-tests were performed by comparing numbers of missing teeth between arches for each analyzed individual. For all the analyses, a p value less than 0.05 was considered statistically significant.

RESULTS

FTA families and novel PAX9 mutations

Whole exome analyses for a cohort of Taiwanese patients with non-syndromic tooth agenesis identified five individuals carrying different disease-causing *PAX9* mutations, all of which have never been reported. The families of these individuals were subsequently recruited and characterized. In all families, the pedigree showed a dominant pattern of disease inheritance, and the mutation segregated with the presence of tooth agenesis (Figure 1). All affected individuals had no anomalies in facial appearance, hair, nails, skin, and heat tolerance.

A total of nine subjects, including seven oligodontia and two hypodontia cases, from five unrelated families were characterized. The pattern of missing teeth for each individual is summarized in Table 1. In general, posterior teeth were more frequently involved than anterior ones. Particularly,

second molars were absent in all individuals, including the two hypodontia cases from Family 3. Maxillary teeth appeared to be more affected than mandibular ones in a given individual. Noticeably, while the Family 2 proband had 17 missing teeth, his brother had only 8, excluding third molars, indicating a highly variable expressivity. Available dental records, including photographs and radiographs, for members of each family are provided in Figures S1 to S6.

The respective *PAX9* mutations identified from each family are NM_006194.4: c.771+4A>G, c.352C>T (p.Pro118Ser), c.247G>T (p.Val83Leu), c.566_588dup (p.Ser197Argfs*23), and c.19G>A (p.Glu7Lys). None of these variants are listed in Genome Aggregation Database (gnomAD)[20] or the Taiwan BioBank database.[21] For the Family 2 proband, a heterozygous splice-site mutation in *WNT10A*, NM_025216.3:c.376+1G>A, was also detected. For the three affected individuals in this family, while the mother and brother also carried the *PAX9* mutation, only the proband had the *WNT10A* mutation inherited from the father, who was reported to have no missing teeth. This segregation pattern of variants suggested a plausible synergistic effect from the *PAX9* and *WNT10A* mutations that caused severe oligodontia in the proband. Similarly, a reported heterozygous *WNT10A* pathogenic mutation c.637G>A (p.Gly213Ser) was identified in the Family 5 proband, indicating a potential contribution to the disease phenotype.

Analyses of PAX9 mutations

We conducted minigene splicing assays to investigate if the *PAX9* c.771+4A>G splice-site mutation alters normal RNA slicing (Figure 2A). While the wild-type construct generated a 404-bp PCR amplicon corresponding to the correctly spliced exon 4, two smaller products were detected from the mutant minigene. The main product (394 bp) came from an mRNA transcript that used an upstream

cryptic splice donor site that shortened dxon 4 by 10 nucleotides, causing a -1 frameshift that terminates prematurely in exon 5. The resulting truncated PAX9 protein, (p.Tyr255Hisfs*30), would contain 29 aberrant amino acids following Lys[254] (Figure 2B). The other amplicon (264 bp) resulted from a transcript that completely skipped exon 4, indicating a loss of splice-site recognition due to the mutation. Deletion of the 140 bps of exon 4 causes a frameshift ending at a premature stop codon in exon 5. This shift into the -2 reading frame truncates PAX9 after Gln[210] and adds 58 extraneous amino acids (p.Val211Glyfs*59). These two aberrantly spliced mRNA transcripts should escape nonsense-mediated decay (NMD) and produce mutant PAX9 proteins. This analysis demonstrated that the c.771+4A>G *PAX9* mutation found in Family 1 is disease causing.

All of the three missense mutations substitute highly-conserved amino acid residues within the paired domain (PD) of PAX9 (Figure 2C) and are predicted to be probably damaging, with a PolyPhen-2 score of 0.999 (p.Glu7Lys), 0.997 (p.Val83Leu), and 1.000 (p.Pro118Ser), respectively.[22] Both SIFT (Sorting Intolerant From Tolerant)[23] and MutationTaster2021[24] also predict these variants to be deleterious, except that p.Glu7Lys is considered tolerated by SIFT but gets a borderline score of 0.057. The 23-bp duplication (c.566_588dup) identified from Family 4 would shift the reading frame at the 3′ part of exon 3 and replace the PAX9 C-terminal 145 amino acids with 22 extraneous ones (p.Ser197Argfs*23). However, the resulting premature termination codon from the frameshift is located at the 5′ end of exon 4, which is likely to subject the mutant transcript to NMD. Therefore, the c.566_588dup *PAX9* mutation is likely to be a null allele that causes tooth agenesis.

Functional effects of *PAX9* missense mutations

To understand the molecular consequences of the identified *PAX9* missense mutations, we first evaluated the intracellular localization of their resulting mutant proteins using immunofluorescence (Figure 3A). Positive signals were detected in the cell nuclei for all the overexpressed PAX9s, except for p.Pro118Ser, that showed immunoreactivity in both nuclei and cytoplasm, indicating a mislocalization of PAX9 caused by this specific mutation. We further analyzed the ability of the PAX9 mutants to induce transcription by using luciferase reporter constructs with two different promoters (Figure 3B). In the reporter driven by an artificial promoter with PAX9 binding sites, the wild-type PAX9 exhibited a significant increase (4.4-fold) in relative luciferase activity (RLA) compared to the empty-vector control, which validated our system for PAX9-induced transcriptional activation. However, the RLAs of p.Glu7Lys, p.Val83Leu, and p.Pro118Ser mutants were only 36.6%, 36.3%, and 32.7% that of wild-type, respectively. This demonstrates not only that PAX9 transactivating ability was significantly compromised by these mutations but that the mutant proteins produced a dominant negative effect on the wild-type PAX9. The *BMP4*-driven reporter showed similar results to those of the artificial promoter except that the p.Glu7Lys mutant exhibited a comparable transactivation ability with the wild-type (Figure 3B).

LEF1 and *AXIN2* are two direct canonical Wnt target genes.[25] To investigate the molecular mechanism underlying the synergistic effect of *PAX9* and *WNT10A* mutations, we tested if Wnt/β-catenin signaling could be modulated by PAX9. When we overexpressed wild-type *PAX9* in dental pulp cells, *LEF1* and *AXIN2* expression was significantly upregulated relative to the empty vector control (Figure 3C). Increased expression was also observed for *BMP4* and *MSX1*, and these effects were partially abolished when the p.Pro118Ser mutated *PAX9* was expressed.

Genotype–phenotype correlations in *PAX9*-associated FTA

A literature review of 37 articles and our current study identified 63 *PAX9* disease-causing mutations and 176 subjects (Figure 4A; Table S1). The genetic defects included 31 missense (M, 49%), 7 nonsense (N, 11%), and 18 frameshift (F, 29%) mutations. Four mutations at the translation initiation codon and 3 variants involving splice sites were categorized as mutations with unknown protein changes (U, 11%) (Figure 4B). Overall, the number of missing teeth (*No.ms*) in each individual varied significantly, ranging from 0 to 28, excluding third molars (mean, 10.23). For each mutation type, the mean *No.ms* was 8.14 (M), 10.24 (N), 12.19 (F), and 12.32 (U) respectively, indicating a relatively milder disease phenotype is caused by *PAX9* missense mutations (Figure 4C; Figure S7). Only 31 individuals (18%) showed a hypodontia phenotype, whereas oligodontia occurred in 82% cases. The majority (81%) of hypodontia patients carried missense mutations (Table S2).

Agenesis mainly involved four posterior teeth: U7 (67–98%), L7 (69–95%), U6 (50–92%), and U5 (45–84%) (Figure 4D). While mandibular second bicuspids (L5) were relatively less affected (37–43%), agenesis of mandibular first molars (L6) varied significantly for different mutation types (20–70%). Other teeth were rarely affected, with less than 30% absence, except for mandibular central incisors (L1, 29–59%). For each individual, more maxillary teeth were missing than mandibular ones, giving a mean difference of 1.31 in *No.ms*. A maxillary tooth was significantly more affected than its mandibular counterpart, except for second molars and central incisors (Figure 4E). While second molars of both arches were comparably absent, L1 was more commonly missing than U1. For hypodontia cases, the pattern of missing teeth varied significantly among individuals, with L7 (35%), U7 (27%), and U5 (24%) being the most frequently affected teeth (Figure S8A).

We further analyzed if the location of missense mutations and the nature of truncation mutations were associated with the disease phenotype. Among 31 missense mutations, only 2 were outside of the PD. The other 29 were located at PAI (17), linker (5), and RED (7) subdomains within the PD (Table S3). While cases with a missense mutation outside the PD appeared to have fewer missing teeth, the pattern of missing teeth among individuals with mutations within different subdomains of the PD was comparable (Figure S8C, D). For nonsense and frameshift mutations, we regrouped them into two categories based upon whether the mutations would likely cause NMD of mutant transcripts. The *No.ms*. and distribution pattern was generally similar between the two groups (Figure S8B).

DISCUSSION

Two previous systematic reviews have shown that mutations in different FTA genes cause distinct patterns of dental agenesis.[9, 11] This study further defines the unique phenotypes caused by *PAX9* mutations. In general, mandibular second bicuspids and maxillary lateral incisors are the most commonly affected teeth in patients with tooth agenesis.[1] However, these two tooth types are not frequently missing in *PAX9*-associated FTA. Instead, agenesis of all second molars and maxillary first molars and second bicuspids appears to be its typical manifestation, which is particularly evident in severe oligodontia cases caused by loss-of-function (N, F, and U) mutations. Human permanent molars without preceding deciduous teeth develop from continual lamina, the distal extension of the dental lamina from second primary molars, through additional tooth formation.[26] The frequent molar agenesis in *PAX9*-associated FTA indicates a critical role for PAX9 in this developmental process of serially adding teeth. Another distinct feature of tooth agenesis caused by *PAX9* mutations is that the maxillary teeth are more commonly involved than the mandibular ones in general and for specific

tooth types, except for second molars and central incisors. This differential involvement suggests that development of maxillary permanent teeth in humans might require more *PAX9* expression than that of mandibular teeth and is therefore more susceptible to *PAX9* mutations. It has been demonstrated that mice carrying different hypomorphic and amorphic *Pax9* alleles exhibited variable severities of tooth agenesis due to differential reduction of *Pax9* gene dosage.[27] However, unlike humans, the lower dentition of these mice was more greatly affected compared to the upper. This phenotypic discrepancy might result from different molecular mechanisms underlying development of primary and secondary dentitions, as mice are monophyodont rather than diphyodont. The dosage effect of the hypomorphic *Pax9* mice also corresponds to our finding in this study that differential disease severity is associated with different types of *PAX9* pathogenic variants, with frameshift and nonsense mutations generally causing more missing teeth than missense variants. This genotype–phenotype correlation not only provides insight into the pathogenesis of *PAX9*-associated FTA but provides a reference for clinical counseling.

In this study, the Family 2 proband, who carried both *PAX9* c.352C>T and *WNT10A* c.376+1G>A mutations, exhibited a much more severe phenotype than his brother, who had only the *PAX9* mutation, suggesting a potential synergistic effect of the two mutations on disease expressivity. This hypothesis was supported by the distinct missing tooth pattern of these two siblings; while the brother had a typical *PAX9* pattern of preserving maxillary laterals and mandibular second bicuspids, the proband had all of these teeth missing. The *WNT10A* c.376+1G>A mutation is a rare variant, with a MAF of 0.0002 for the East Asian (EAS) population[20] and has been shown to cause mild hypodontia or no tooth agenesis in heterozygous carriers.[28, 29] However, when combined with a missense *PAX9* mutation, which causes an average of ~8 missing teeth, it leads to a full-blown phenotype of 17 missing teeth in our patient. This genetic synergism was also partly substantiated by the dental

phenotype of the Family 5 proband, who had *PAX9* c.19G>A and *WNT10A* c.637G>A mutations, and a missing tooth pattern different from that of *PAX9*-associated FTA we characterized here. The *WNT10A* c.637G>A variant is a missense mutation, p.Gly213Ser, that has been identified in many FTA cases[30, 31] and shown to abolish WNT transcriptional activity of wild-type WNT10A *in vitro*.[32] As the mutation is a relatively common variant in East Asian (MAF = ~0.028),[20] it could serve as a significant genetic modifier that causes variable expressivity in FTA cases of the EAS population. Therefore, when analyzing the genetic etiology of FTA, it is critical to conduct a comprehensive mutational analysis of all FTA candidate and odontogenesis-related genes. However, during our literature review, we noticed that the vast majority (33/37) of studies used targeted gene approaches to identify *PAX9* mutations, which might fail to find other sequence variants that potentially contribute to the disease phenotype. Our cases demonstrated the importance of a comprehensive approach, primary through whole exome/genome sequencing, to make a more accurate genetic diagnosis for FTA patients in the clinic. Furthermore, the genetic synergy we demonstrated here not only provides plausible explanations for clinical phenotypes but indicates a genetic interaction between PAX9 and Wnt signaling during tooth development. In early odontogenesis, *Pax9* was specifically expressed by dental mesenchyme and is critical for *Bmp4* expression.[27, 33] On the other hand, Wnt/β-catenin signaling is active primarily at the enamel knot and adjacent mesenchyme.[34] While *Pax9* expression does not spatially correspond to that of Wnt/β-catenin activity, *Pax9* null mice have been shown to have a significantly reduced Wnt signaling activity in dental mesenchyme.[33] In contrast, in the absence of mesenchymal β-catenin, expression of *Pax9* is not altered, suggesting an epistatic role for *Pax9* in mesenchymal Wnt activity.[35] Here, we demonstrate that *PAX9* overexpression can upregulate both *LEF1* and *AXIN2* expression, further supporting this regulatory role for *PAX9* in Wnt signaling during odontogenesis. Nevertheless, it was recently demonstrated that activation of Wnt/β-catenin

signaling rescues palatogenesis, but not odontogenesis, in *Pax9*-deficient mice,[36] suggesting that canonical Wnt signaling is probably not a direct downstream effector of PAX9 during tooth formation. Further investigation is warranted to discern the PAX9–Wnt signaling interaction in tooth development.

ACKNOWLEDGMENTS

We thank the families for participation in this study and Stuart Tompson at the University of Wisconsin-Madison for his generous gift of pSPL3 plasmid. This study was supported by Ministry of Science and Technology in Taiwan (MOST) Grants 108-2314-B-002-038-MY3 (S.-K.W.) and 111-2314-B-002-111-MY3 (S.-K.W.); National Taiwan University Hospital (NTUH) Grants 110-N4809 (S.-K.W.) and 110-S4947 (Y.-L.W.); and National Institutes of Health Grants R01DE027675 (J.P.S.) and R56DE015846 (J.C.-C.H.).

COMPETING INTERESTS

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

K.-Y.C. contributed to conception and design, data acquisition and analysis, helped draft and critically revised the manuscript. Y.-L.W., J.-T.C., and C.-H.L. contributed to design, data acquisition and analysis, and critically revised the manuscript. C.-C.J.Y., Y.-J.C., and H.-W.C. contributed to design,

data acquisition, and critically revised the manuscript. J.P.S. and J.C.-C.H. contributed to conception and design, data analysis and interpretation, and critically revised the manuscript. S.-K.W. recruited the subjects, contributed to conception and design, data acquisition and analysis, helped draft and critically revised the manuscript.

References

1. Nieminen, P. 2009. Genetic basis of tooth agenesis. *J Exp Zool B Mol Dev Evol*. **312b**: 320-342.
2. Juuri, E. & A. Balic. 2017. The Biology Underlying Abnormalities of Tooth Number in Humans. *J Dent Res*. **96**: 1248-1256.
3. Vastardis, H., N. Karimbux, S.W. Guthua*, et al.* 1996. A human MSX1 homeodomain missense mutation causes selective tooth agenesis. *Nat Genet*. **13**: 417-421.
4. Stockton, D.W., P. Das, M. Goldenberg*, et al.* 2000. Mutation of PAX9 is associated with oligodontia. *Nat Genet*. **24**: 18-19.
5. Lammi, L., S. Arte, M. Somer*, et al.* 2004. Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. *Am J Hum Genet*. **74**: 1043-1050.
6. Song, S., D. Han, H. Qu*, et al.* 2009. EDA gene mutations underlie non-syndromic oligodontia. *J Dent Res*. **88**: 126-131.
7. Bohring, A., T. Stamm, C. Spaich*, et al.* 2009. WNT10A mutations are a frequent cause of a broad spectrum of ectodermal dysplasias with sex-biased manifestation pattern in heterozygotes. *Am J Hum Genet*. **85**: 97-105.
8. Massink, M.P., M.A. Créton, F. Spanevello*, et al.* 2015. Loss-of-Function Mutations in the WNT Co-receptor LRP6 Cause Autosomal-Dominant Oligodontia. *Am J Hum Genet*. **97**: 621-626.
9. Fournier, B.P., M.H. Bruneau, S. Toupenay*, et al.* 2018. Patterns of Dental Agenesis Highlight the Nature of the Causative Mutated Genes. *J Dent Res*. **97**: 1306-1316.

10. Kim, J.W., J.P. Simmer, B.P. Lin, *et al.* 2006. Novel MSX1 frameshift causes autosomal-dominant oligodontia. *J Dent Res*. **85**: 267-271.

11. Zhou, M., H. Zhang, H. Camhi, *et al.* 2021. Analyses of oligodontia phenotypes and genetic etiologies. *Int J Oral Sci*. **13**: 32.

12. Wong, S.W., D. Han, H. Zhang, *et al.* 2018. Nine Novel PAX9 Mutations and a Distinct Tooth Agenesis Genotype-Phenotype. *J Dent Res*. **97**: 155-162.

13. He, H., D. Han, H. Feng, *et al.* 2013. Involvement of and interaction between WNT10A and EDA mutations in tooth agenesis cases in the Chinese population. *PLoS One*. **8**: e80393.

14. Chu, K.Y., Y.L. Wang, Y.R. Chou, *et al.* 2021. Synergistic Mutations of LRP6 and WNT10A in Familial Tooth Agenesis. *J Pers Med*. **11**.

15. Pérez-Pérez, J.M., H. Candela & J.L. Micol. 2009. Understanding synergy in genetic interactions. *Trends Genet*. **25**: 368-376.

16. Arte, S., S. Parmanen, S. Pirinen, *et al.* 2013. Candidate gene analysis of tooth agenesis identifies novel mutations in six genes and suggests significant role for WNT and EDA signaling and allele combinations. *PLoS One*. **8**: e73705.

17. Tompson, S.W. & T.L. Young. 2017. Assaying the Effects of Splice Site Variants by Exon Trapping in a Mammalian Cell Line. *Bio Protoc*. **7**.

18. Czerny, T., G. Schaffner & M. Busslinger. 1993. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev*. **7**: 2048-2061.

19. Chen, J.T., C.H. Lin, H.W. Huang, *et al.* 2021. Novel REST Truncation Mutations Causing Hereditary Gingival Fibromatosis. *J Dent Res*. **100**: 868-874.

20. Karczewski, K.J., L.C. Francioli, G. Tiao, *et al.* 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. **581**: 434-443.

21. Chen, C.H., J.H. Yang, C.W.K. Chiang, *et al.* 2016. Population structure of Han Chinese in the modern Taiwanese population based on 10,000 participants in the Taiwan Biobank project. *Hum Mol Genet*. **25**: 5321-5331.

22. Adzhubei, I.A., S. Schmidt, L. Peshkin, *et al.* 2010. A method and server for predicting damaging missense mutations. *Nat Methods*. **7**: 248-249.

23. Sim, N.L., P. Kumar, J. Hu, *et al.* 2012. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res*. **40**: W452-457.

24. Steinhaus, R., S. Proft, M. Schuelke, *et al.* 2021. MutationTaster2021. *Nucleic Acids Res*. **49**: W446-w451.

25. Jho, E.H., T. Zhang, C. Domon, *et al.* 2002. Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. *Mol Cell Biol*. **22**: 1172-1183.

26. Hovorakova, M., H. Lesot, M. Peterka, *et al.* 2018. Early development of the human dentition revisited. *J Anat*. **233**: 135-145.

27. Kist, R., M. Watson, X. Wang, *et al.* 2005. Reduction of Pax9 gene dosage in an allelic series of mouse mutants causes hypodontia and oligodontia. *Hum Mol Genet*. **14**: 3605-3617.

28. Kantaputra, P., M. Kaewgahya, D. Jotikasthira, *et al.* 2014. Tricho-odonto-onycho-dermal dysplasia and WNT10A mutations. *Am J Med Genet A*. **164a**: 1041-1048.

29. Tardieu, C., S. Jung, K. Niederreither, *et al.* 2017. Dental and extra-oral clinical features in 41 patients with WNT10A gene mutations: A multicentric genotype-phenotype study. *Clin Genet*. **92**: 477-486.

30. Yang, J., S.K. Wang, M. Choi, *et al.* 2015. Taurodontism, variations in tooth number, and misshapened crowns in Wnt10a null mice and human kindreds. *Mol Genet Genomic Med*. **3**: 40-58.

31. Song, S., R. Zhao, H. He, *et al.* 2014. WNT10A variants are associated with non-syndromic tooth agenesis in the general population. *Hum Genet*. **133**: 117-124.

32. Zeng, Y., E. Baugh, S. Akyalcin, *et al.* 2021. Functional Effects of WNT10A Rare Variants Associated with Tooth Agenesis. *J Dent Res*. **100**: 302-309.

33. Peters, H., A. Neubüser, K. Kratochwil, *et al.* 1998. Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. *Genes Dev*. **12**: 2735-2747.

34. Liu, F., E.Y. Chu, B. Watt, *et al.* 2008. Wnt/beta-catenin signaling directs multiple stages of tooth morphogenesis. *Dev Biol*. **313**: 210-224.

35. Chen, J., Y. Lan, J.A. Baek, *et al.* 2009. Wnt/beta-catenin signaling plays an essential role in activation of odontogenic mesenchyme during early tooth development. *Dev Biol*. **334**: 174-185.

36. Li, C., Y. Lan, R. Krumlauf, *et al.* 2017. Modulating Wnt Signaling Rescues Palate Morphogenesis in Pax9 Mutant Mice. *J Dent Res*. **96**: 1273-1281.

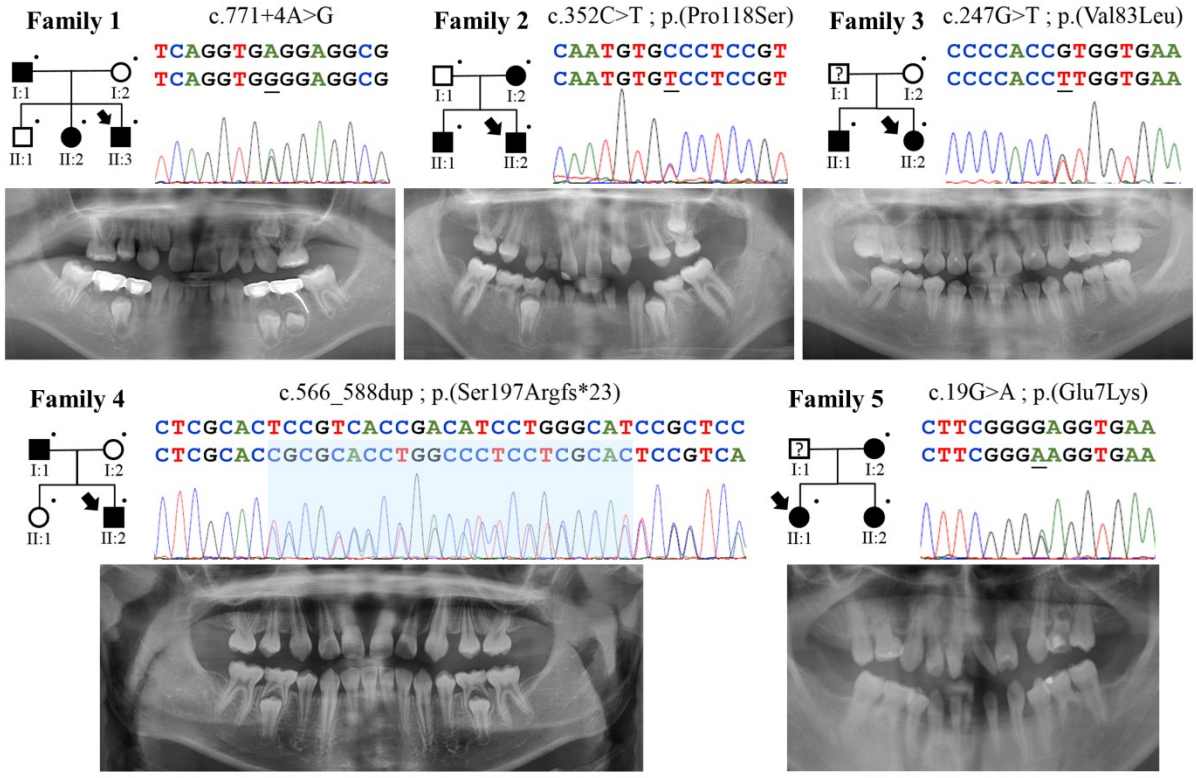


Figure 1. Five FTA families with novel *PAX9* disease-causing mutations. All family pedigrees indicate a dominant pattern of disease inheritance. The dots mark subjects who provided DNA samples for genetic testing. The panoramic radiographs from each proband of the 5 families show respectively 11, 17, 4, 8, and 12 congenitally missing teeth, excluding third molars. All *PAX9* mutations are single-nucleotide substitutions except for a 23-bp duplication (c.566_588dup) from Family 4. The duplicated nucleotides are highlighted in light blue. All mutations are designated based on the *PAX9* cDNA reference sequence NM_006194.4 and protein reference sequence NP_006185.1.

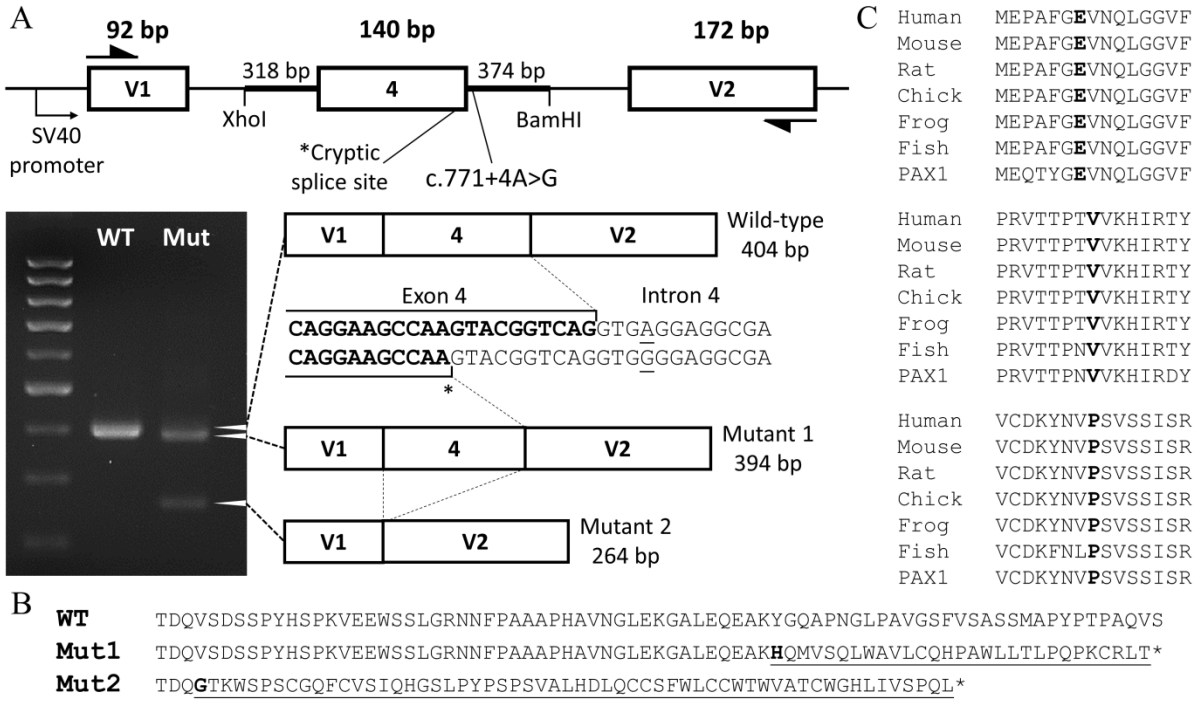


Figure 2. Analyses of *PAX9* disease-causing mutations. (A) The wild-type (WT) and mutant (Mut) minigene constructs were generated by cloning human *PAX9* exon 4 and its flanking sequences into the pSPL3 vector. The agarose gel image shows a single RT-PCR amplification product (404 bp) in WT and two mutant amplicons (394 and 264 bp) from the mutant minigene. While the larger product is caused by a cryptic splice site within exon 4 (asterisk), the smaller one is caused by exon 4

skipping. (B) Alignment of wild-type PAX9 amino acid sequence (Glu223–Ser284) with those of truncated PAX9s that are hypothetically generated from the two mutant transcripts: p.(Tyr255Hisfs*30) and p.(Val211Glyfs*59). (C) Alignment of the human PAX1 amino acid sequence with PAX9 protein sequences from human (*Homo sapiens*), mouse (*Mus musculus*), rat (*Rattus norvegicus*), chick (*Gallus gallus*), frog (*Xenopus tropicalis*), and fish (*Danio rerio*). The substituted amino acids in p.Glu7Lys (top), p.Val83Leu (middle), and p.Pro118Ser (bottom) are in bold.

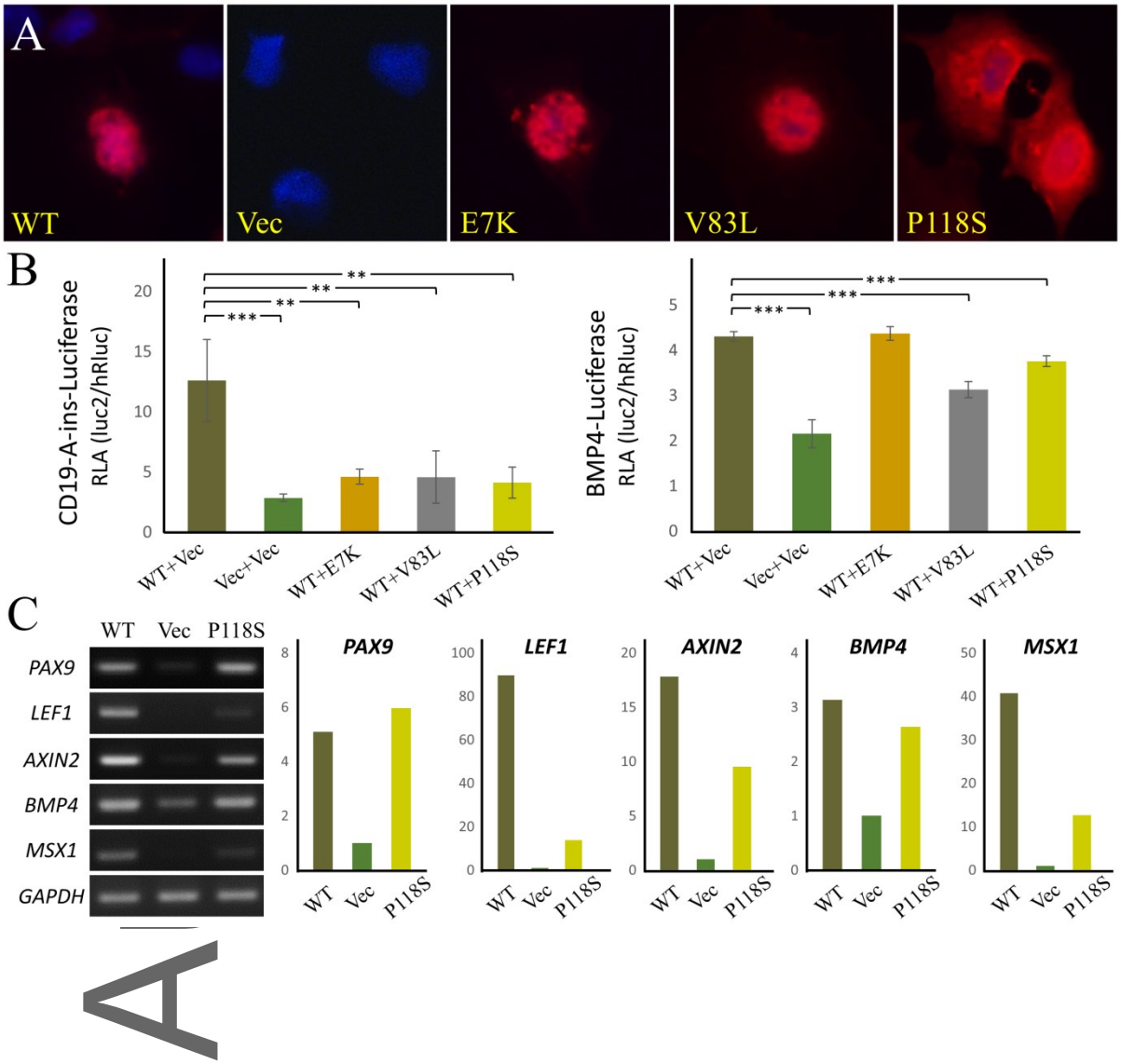
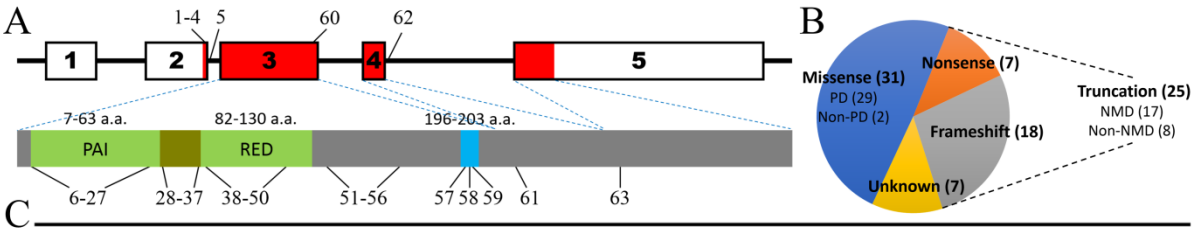
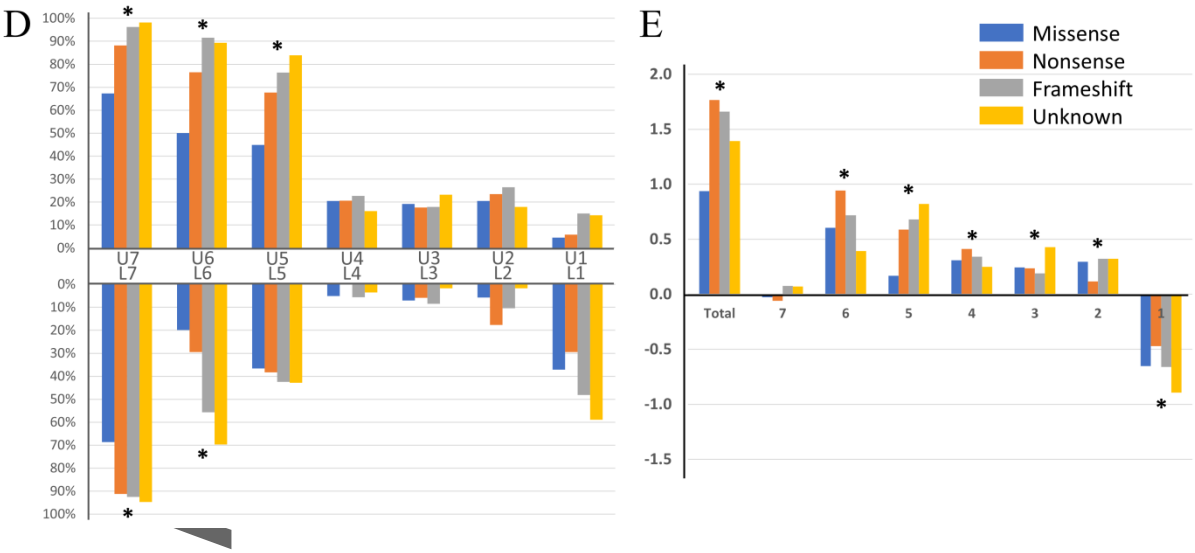


Figure 3. Molecular characterization of the missense mutations. (A) Anti-c-Myc immunostaining (red) of wild-type and mutant PAX9 proteins overexpressed in COS7 cells. Nuclei are stained blue with DAPI. (B) Relative luciferase activity (RLA) in dual luciferase assays with co-expression of wild-type and mutated *PAX9* constructs or the empty vector. The left chart shows the results of using the CD19-A-ins-Luciferase reporter, while the right chart demonstrates the RLA of using the BMP4-Luciferase construct. (C) The agarose gel images on the left show RT-PCR amplification products of indicated genes using RNA from transfected dental pulp cells. The charts on the right present quantification of corresponding gene bands on the gel images using ImageJ. The numbers on each chart are normalized to that of empty vector control. WT, wild-type; Vec, empty vector; E7K, p.Glu7Lys; V83L, p.Val83Leu; P118S, p.Pro118Ser; $^{**}p < 0.05$; $^{***}p < 0.01$.



C

Mutation type	Mutation No.	Case No.	Range (missing tooth No.)	Mean (missing tooth No.)	*p*-value	Hypodontia cases
Missense (M)	31	78	23 (0~22)	8.14	-	25
Nonsense (N)	7	17	19 (2~20)	10.24	0.354	3
Frameshift (F)	18	53	27 (2~28)	12.19	0.000*	2
Unknown (U)	7	28	16 (4~19)	12.32	0.001*	1



Subject	Age		8	7	6	5	4	3	2	1	1	2	3	4	5	6	7	8	No	*PAX9* mutation	*WNT10A* mutation

Figure 4. Genotype–phenotype correlations in *PAX9*-associated FTA. (A) Location of 59 reported *PAX9* disease-causing mutations on gene (top) and protein (bottom) structure diagrams. The numbering of each mutation corresponds to Table S1. Exons are numbered boxes. The coding region is marked in red. The PAX9 protein domain structure is based upon annotations of P55771 in the UniProt database. The paired domain (PD) contains 3 subdomains: PAI (green), linker (dark yellow), and RED (green). The light blue box represents the octapeptide motif. (B) The pie chart shows the number of mutations in each category. Only two missense mutations are located outside of PD. Truncation mutations, including nonsense and frameshift mutations, are further grouped based upon if they likely undergo nonsense mediated decay (NMD). (C) Summary table of descriptive statistics of missing tooth numbers in each category. (D) Correlations between the percentage of missing teeth at each tooth position and the mutation category. (E) Correlations between the differential (subtractive) missing tooth number of upper (U) and lower (L) jaws at each tooth position and the mutation category. a.a., amino acid; No., number; $*p < 0.05$.

Table 1. Pattern of missing teeth in affected individuals from each FTA family.

Family 1	I:1	48	Max	X	X	X			X					X		X	X	X	X	13	c.771+4A>G p.(?)	-
			Man	X	X	X					X	X					X	X	X			
	II:2	16	Max	X	X	X	X							X		X	X	X	X	13		-
			Man	X	X	X					X	X					X	X	X			
	II:3	8	Max	?	X	X	X									X	X	X	?	11		-
			Man	?	X		X				X	X						X	?			
Family 2	II:1	14	Max	X	X	X			X					X			X	X	X	8	c.352C>T p.(P118S)	-
			Man	X	X													X	X			
	II:2	11	Max	X	X	X	X		X	X			X				X	X	X	17		c.376+1G>A
			Man	X	X		X		X	X	X	X	X			X		X	X			
Family 3	II:1	21	Max		X													X	X	5	c.247G>T p.(V83L)	-
			Man	X	X					X								X	X			
	II:2	17	Max	X	X													X	X	4		-
			Man	X	X													X	X			
Family 4	II:2	9	Max	X	X	X	X									X	X	X	X	8	c.566_588dup p.(S197Rfs*23)	-
			Man	X	X													X	X			
Family 5	II:1	28	Max	X	X		X	X							X	X		X	X	12	c.19G>A p.(E7K)	c.637G>A p.(G213S)

Note: The symbol for each subject refers to the family pedigrees in Figure 1. The subject's age is that at the time of recruitment. The number (No.) of total missing teeth excludes third molars. Max, maxillary; Man, mandibular.