Identification of Novel Candidate Genes Associated with Non-Syndromic Tooth Agenesis in Mongolian Families

Dejidnorov Semjid

The Graduate School

Yonsei University

Department of Dentistry

Identification of Novel Candidate Genes Associated with Non-Syndromic Tooth Agenesis in Mongolian Families

A Dissertation

Submitted to the Department of Dentistry

and the Graduate School of Yonsei University

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy in Dental Science

Dejidnorov Semjid

December 2023

This certifies that the Doctoral Dissertation

of Dejidnorov Semjid is approved.

Thesis Supervisor: Jae Hoon Lee

Jee-Hwan Kim

Sung-Won Cho

Je Seon Song

Sanguk Kim

The Graduate School

Yonsei University

December 2023

Acknowledgement

First and foremost, I thank my God who created Heaven and Earth, for giving me the strength and motivation I needed during the challenging moments of completing this dissertation. Thank you for reminding me that you are always with me.

I would like to express my profound gratitude to the Republic of Korea for providing me with the invaluable opportunity to pursue my doctoral studies in this enchanting country.

My heartfelt thanks go to Yonsei university, School of Dentistry, Department of Prosthodontics, for their support and abundant resources throughout my doctoral journey. The academic environment and world-class faculty at Yonsei university have played a pivotal role in shaping my research and fostering my personal growth.

I am deeply indebted to my doctoral thesis supervisor, Professor Jae Hoon Lee. Initially, I had doubts about conducting research in this field, but his guidance enabled me to conduct my research successfully. His expertise and mentorship have significantly shaped my research and academic growth as a researcher and as a dentist. I am deeply grateful for his support, and it is an honor to have been his student.

I extend my heartfelt appreciation to my dissertation committee members: Professor Jee-Hwan Kim, Professor Sung-Won Cho, Professor Je Seon Song, and Professor Sanguk Kim for all their meticulous guidance, insightful comments, thought-provoking questions, and constructive feedback throughout the dissertation process.

I am also immensely thankful to Professor Sanguk Kim and Hyung Soo Ahn for their expertise in bioinformatic analysis. I will always appreciate their support and guidance.

I would like to express my gratitude my dear Professor Sapaar Bayarmagnai and Munkhjargal Gantumur, as well as the professors at the Mongolian National University of Medical Sciences, School of Dentistry, Department of Prosthodontics. You have been pivotal in guiding and inspiring me throughout my academic journey, from the very beginning.

A special thanks goes out to my dear friend whose support and friendship have provided me with strength and encouragement.

Last but certainly not least, I express my deepest gratitude to my family for their unwavering support, love, and understanding throughout this challenging yet rewarding journey. Their encouragement and belief in my capabilities have been a constant source of strength, and I am truly fortunate to have them by my side.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Identification of Novel Candidate Genes Associated with Non-Syndromic Tooth Agenesis in Mongolian Families

Dejidnorov Semjid

Department of Dentistry, Graduate School, Yonsei University
(*Directed by Professor Jae Hoon Lee, D.D.S., M.D., Ph.D.)*

Tooth agenesis is a prevalent developmental anomaly characterized by the absence of teeth. The identification of numerous genes linked to the etiology of tooth agenesis contributes to enhancing understanding of tooth development. However, there is still a need fully elucidate novel causal genes and their associated signaling pathways.

This study aimed to identify genetic variants associated with non-syndromic TA in nine families from Mongolia using whole-exome sequencing (WES) and bioinformatic analysis. The study enrolled 41 participants, including three inherited and six non-inherited families. WES analysis was performed on 14 saliva samples from individuals with non-syndromic TA.

The potential candidate genes were identified through variant filtering and segregation analysis. The filtered variants were then analyzed in silico mutation impact analysis. WES

analysis identified 21 variants associated with TA, and 5 of these variants met all filtering criteria. These variants were located in the exon region of MAST4, ITGA6, PITX2, CACNA1S, and CDON genes.

The variant in PITX2 was found in eight participants from both inherited and non-inherited families, while the MAST4 variant was identified in six participants from inherited families. Additionally, the CDON variant was identified in five participants from both inherited and non-inherited families. Gene set enrichment analysis revealed the significance of five candidate genes associated with functional terms related to focal adhesion and the calcium channel complex.

The study has identified novel candidate gene variants associated with non-syndromic TA in Mongolian families. Furthermore, the results emphasize the significance of genes related to the Wnt signaling pathway, focal adhesion and calcium channel complexes in the genetic mechanisms of tooth development.

Keywords: Tooth agenesis, Genetic variants, Whole-exome sequencing, Bioinformatic analysis, In silico mutation, Mongolian population

Identification of Novel Candidate Genes Associated with Non-Syndromic Tooth Agenesis in Mongolian Families

Department of Dentistry,
Graduate School, Yonsei University
(Directed by Professor Jae Hoon Lee, D.D.S., M.D., Ph.D.)

Dejidnorov Semjid

I. INTRODUCTION

Tooth agenesis (TA) is a common developmental anomaly that results in the congenital absence of one or more permanent teeth in humans. It affects approximately 200 million individuals worldwide, and its incidence varies by geography, population, and race [1, 2]. TA can lead to several complications, including oral symptoms, masticatory dysfunction, physiological issues, speech impairments, aesthetic concerns, and financial burdens, which can significantly affect the quality of life [3, 4].

Tooth development begins with sequential and reciprocal interactions between embryonic tissues, such as oral epithelium and ectomesenchyme [5]. This process is regulated by various signaling pathways, including BMP, FGF, SHH, TNF, and WNT, which are essential for the

tooth bud, cap, and bell stages [6, 7]. Disruptions in these signaling pathways during tooth development may cause TA [8].

TA can be categorized based on the number of missing permanent teeth, typically excluding third molars. Hypodontia refers to the absence of fewer than six permanent teeth, while oligodontia is defined as the absence of six or more permanent teeth. Anodontia describes the complete absence of all permanent teeth [9] . The worldwide prevalence of third molar agenesis is estimated to range from 5.32% to 56% [10]. Therefore, it is typically excluded from dental numbering due to their common absence. In the majority of cases (83%) involving TA only one or two permanent teeth are affected, while the occurrence of the absence of six permanent teeth is exceptionally rare (0.14%) [11].

TA can also be classified as non-syndromic or syndromic, depending on the involvement of other organs or tissues. Non-syndromic TA, also known as isolated TA, is characterized by dental abnormalities without other tissue symptoms. On the other hand, syndromic TA involves missing teeth and other developmental abnormalities such as nails, hair, skin and sweat glands. Some examples of syndromic TA include cleft lip/plate [12], ectodermal dysplasia [13], Axenfeld-Rieger syndrome, and Witkop syndrome [14-16]. Non-syndromic TA occurs more frequently than syndromic TA.

The etiology of TA can be attributed to genetic or environmental factors [17-20]. However, genetic factors play a more significant role in the pathogenesis of TA [21, 22]. Identifying the genetic factors associated with TA can improve diagnosis and treatment. Recent studies have identified several genes associated with TA, including MSX1 [23], PAX9 [19], BMP4 [24], AXIN2 [25], EDA [26], EDAR [27], EDARADD [28], WNT10A [4], WNT10B [29], LRP6 [30], PITX2 [31], FGFR2 [3], and CACNA1S [3]. Current studies employing WES and bioinformatic analysis have explored TA. Researchers suggest that EDA, EDAR, WNT10A, and PCNT are linked to syndromic TA. Conversely, studies have indicated that genes such as PAX9, MSX1, LRP6, WNT10A, WNT10B, AXIN2, PITX2, and BMP4 are associated with

non-syndromic TA [32-34]. These genes exhibit autosomal-dominant, autosomal-recessive, or X-linked mechanisms of inheritance [35].

The highest genetic risk for non-syndromic TA is attributed to mutated genes that encode components within the canonical Wnt/β-catenin signaling pathway. The Wnt signaling pathway consists of several key components, including extracellular secreted glycoproteins (comprising 19 Wnt ligands in humans), receptors with seven transmembrane spans (Frizzled and LRP5/6), cytoplasmic proteins (DVL, APC, AXIN, GSK3 β , β -catenin), nuclear transcription factors (TCF/LEF), and various other molecules associated with Wnt (such as DKK1, KREMEN 1, ANTXR1 and MSX1) [33].

The initial connection between TA and Wnt pathway became evident when a mutation in the AXIN2 gene was discovered in a family with oligodontia. AXIN2 is responsible for producing a protein that inhibits the Wnt/β-catenin signaling pathway and is prominently expressed in enamel knot and mesenchymal odontoblasts during tooth development [18]. WNT10A, belonging to the Wnt family present in the dental epithelium, interacts with the Frizzled transmembrane receptor and LRP6. This interaction subsequently triggers the activation of Wnt/β-catenin pathway, which leads to abnormal tooth development [33].

MSX1 and PAX9 are among the first genes associated with TA, and their protein products act as transcription factors essential for the tooth germ's development from mesenchymal cells [36, 37]. Mutations in the MSX1 gene, which encodes a transcriptional repressor involved in both the Wnt and BMP4 pathways, have been consistently identified as the cause of non-syndromic TA [38]. Recently, it has been shown that PITX2 contributes to TA and is involved in the early stages of tooth development, including the formation of tooth germs and bud morphology [39].

In recent years, the next-generation sequencing (NGS) approach, specifically WES platforms, such as Illumina, has been extensively used for identifying biomarkers of genetic disease diagnosis [40]. Although only 2% of the human genome consists of exons, 85% of the

genetic variations responsible for highly penetrant diseases reside in this small genome region [31]. WES has improved diagnostic accuracy, shortened the diagnostic process, and is more cost-effective than traditional methods. This approach identifies genetic variations that may contribute to the development of TA. Bioinformatic analysis can then be used to analyze the WES data [41]. It can help identify potential candidate genes and variants that may contribute to the condition, especially in a limited number of participants. Furthermore, it can help predict the functional impact of genetic variants and provide insights into the pathogenesis of the condition.

The study aims to investigate the genetic variants of non-syndromic TA in Mongolian families through the utilization of WES and bioinformatic analysis.

II. MATERIALS AND METHODS

1. Ethics approval

This study was approved by Institutional Review Boards at Ministry of Health in Mongolia (IRB No. 285, Approved on 7th September 2022). This clinical study was conducted in accordance with Helsinki Declaration. All enrolled participants were informed about consent and ascertained. Nine participants were juvenile's, written informed consent was obtained from their guardians.

2. Subjects

A total of 41 individuals from nine Mongolian families were enrolled in the study. Among these individuals, 15 were identified with non-syndromic TA. However, in Family 1, one participant (I-2) who had non-syndromic TA was unable to provide saliva samples for analysis. Consequently, 14 participants who had both clinical data and saliva samples available underwent WES analysis. Participants with TA were identified through clinical examination and panoramic radiograph. It was confirmed that the missing permanent teeth were not due to extraction or injuries. The general clinical examination confirmed that all study participants exhibited non-syndromic TA and had normal hair, skin, sweat glands, facial features and nails.

3. Radiographic assessments

Panoramic radiographs were obtained from 14 participants.

4. Sample collection

Saliva samples were collected from each participant using the Oragene DNA Self-Collection Kit (DNA Genotek Inc., Ottawa, Canada). A 2 ml saliva was then collected and subsequently mixed with a DNA-preserving solution in a tube, following the guidelines set forth by the manufacturer.

5. Control group

The study utilized exome sequencing data from a randomized subsample of 100 healthy individuals (Koreans), obtained from the Ansan-Ansung population consisting of 3703 individuals. These subsamples were provided by the Korea BioBank, Center for Genome Science, National Institute of Health, Korea Centers for Disease Control and Prevention. The individuals selected from the reference population were chosen randomly, regardless of sex and age, and had no history of significant diseases.

6. Whole-exome sequencing

DNA samples were processed for WES using the SureSelectXT Human All Exon V5 kit and Novaseq 6000 platform. The DNA quality was assessed through 1% agarose gel electrophoresis and PicoGreen® dsDNA Assay. The library was prepared by fragmenting genomic DNA, ligating sequencing adapters, and amplifying the adapter-ligated DNA using PCR. A hybridization buffer was then prepared by mixing SureSelect hyb #1, #2, #3 and #4 reagents, and the amplified DNA fragments was concentrated and SureSelect blocks #1, #2, and #3 reagents were added. The DNA blocking agent mixture and hybridization buffer were incubated, and a RNase block was added to the SureSelect oligo capture library, which was then incubated. After adding the hybridization buffer and DNA blocking agent mix to the capture library, the mixture was incubated at 65℃ for 24 hours. The captured library was then washed with SureSelect binding buffer and eluted with nuclease-free water. Finally, the library was amplified and tagged with index tags. The libraries were pooled in equimolar amounts and subjected to sequencing using the Illumina Novaseq 6000 system following the protocol for 2x100 sequencing.

7. Family-level statistical analysis of WES data and variant filtering

The study grouped patients from the same family into one subgroup to enhance the effects of a family inheritance. If not, each family's sole patients are placed in the opposite grouping. The WES data were used to identify variants with only protein-coding transcripts being taken into account.

Fisher's exact test was used to compare the allele frequencies of each mutation between the control and each subgroup. To qualify risky variants of TA, variants with association P value ($p < 0.05$) and Odds ratio (ORs > 1) were considered statistically significant (Table 2). The threshold of OR was set as > 1, which just implies that the variant ratio was higher in the patient subgroup than in control cohort. Then, the significant associations of the variant were evaluated with p-values.

After statistical analysis, the study subsequently collected variants that can be associated with TA by disrupting protein function. To reduce the systematic differences between the sample subgroups, variants on TA-related genes filtered using public databases of gene-phenotype relationships. Additionally, variants related to TA were filtered using associated gene lists from Open Targets (OT) [42] (EFO_0005410, tooth agenesis), Gene Ontology (GO) [43] (GO:0042476, odontogenesis), and Human Phenotype Ontology (HPO) [44] (HP:0000677, oligodontia). This step helps narrow the range of variants to the specific area of interest, reducing the noise signal from the other variants.

Variants with annotation of HIGH impact in SnpEff [45] and variants with MODERATE impact (Table 1) but are predicted to damage its function with the results of computational prediction tools (PolyPhen2 [46] and SIFT [47]) or lower minor allele frequency in the Asian population (ASN MAF [48]) than 0.1 were filtered. Figure 1 flow chart provides a step-by-step visualization of the variant filtration process, providing clarity and transparency regarding the analytical pipeline.

Table 1. Impact category of effects in SnpEff.

Putative Impact	Effects(Sequence Ontology term)
HIGH	chromosome_number_variation
HIGH	exon_loss_variant
HIGH	frameshift_variant
HIGH	rare_amino_acid_variant
HIGH	splice_acceptor_variant
HIGH	splice_donor_variant
HIGH	start_lost
HIGH	stop_gained
HIGH	stop_lost
HIGH	transcript_ablation
MODERATE	3_prime_UTR_truncation+exon_loss
MODERATE	5_prime_UTR_truncation+exon_loss_variant
MODERATE	coding_sequence_variant
MODERATE	disruptive_inframe_deletion
MODERATE	disruptive_inframe_insertion
MODERATE	inframe_deletion
MODERATE	inframe_insertion
MODERATE	missense_variant
MODERATE	regulatory_region_ablation
MODERATE	splice_region_variant
MODERATE	TFBS_ablation
LOW	5_prime_UTR_premature start_codon_gain_variant
LOW	initiator_codon_variant
LOW	splice_region_variant
LOW	start_retained

LOW	stop_retained_variant
LOW	synonymous_variant
MODIFIER	3_prime_UTR_variant
MODIFIER	5_prime_UTR_variant
MODIFIER	coding_sequence_variant
MODIFIER	conserved_intergenic_variant
MODIFIER	conserved_intron_variant
MODIFIER	downstream_gene_variant
MODIFIER	exon_variant
MODIFIER	feature_elongation
MODIFIER	feature_truncation
MODIFIER	gene_variant
MODIFIER	intergenic_region
MODIFIER	intragenic_variant
MODIFIER	intron_variant
MODIFIER	mature_miRNA_variant
MODIFIER	miRNA
MODIFIER	NMD_transcript_variant
MODIFIER	non_coding_transcript_exon_variant
MODIFIER	non_coding_transcript_variant
MODIFIER	regulatory_region_amplification
MODIFIER	regulatory_region_variant
MODIFIER	TF_binding_site_variant
MODIFIER	TFBS_amplification
MODIFIER	transcript_amplification
MODIFIER	transcript_variant
MODIFIER	upstream_gene_variant

Table 2. The threshold of Odds ratio (ORs).

Subgroup	CHROM	ID	REF	ALT	Effect	Impact	Gene_Name	OR	Fisher_p
Family 1	2	.	G	GA	frameshift_variant	HIGH	ITGA6	99	0.00171572
	5	rs201910335	C	CGCT	disruptive_inframe_insertion	MODERATE	MAST4	64	0.00395857
Family 2	5	rs201910335	C	CGCT	disruptive_inframe_insertion	MODERATE	MAST4	32	0.00941702
	4	rs2278782	G	A	stop_gained	HIGH	PITX2	6.33333333	0.04701823
Family 3	1	rs3850625	G	A	missense_variant	MODERATE	CACNA1S	27.5714286	0.00154489
	5	rs201910335	C	CGCT	disruptive_inframe_insertion	MODERATE	MAST4	32	0.00941702
Non-Fam	4	rs2278782	G	A	stop_gained	HIGH	PITX2	4.52380952	0.02661926
	11	rs12274923	G	A	missense_variant	MODERATE	CDON	4.5	0.03387269

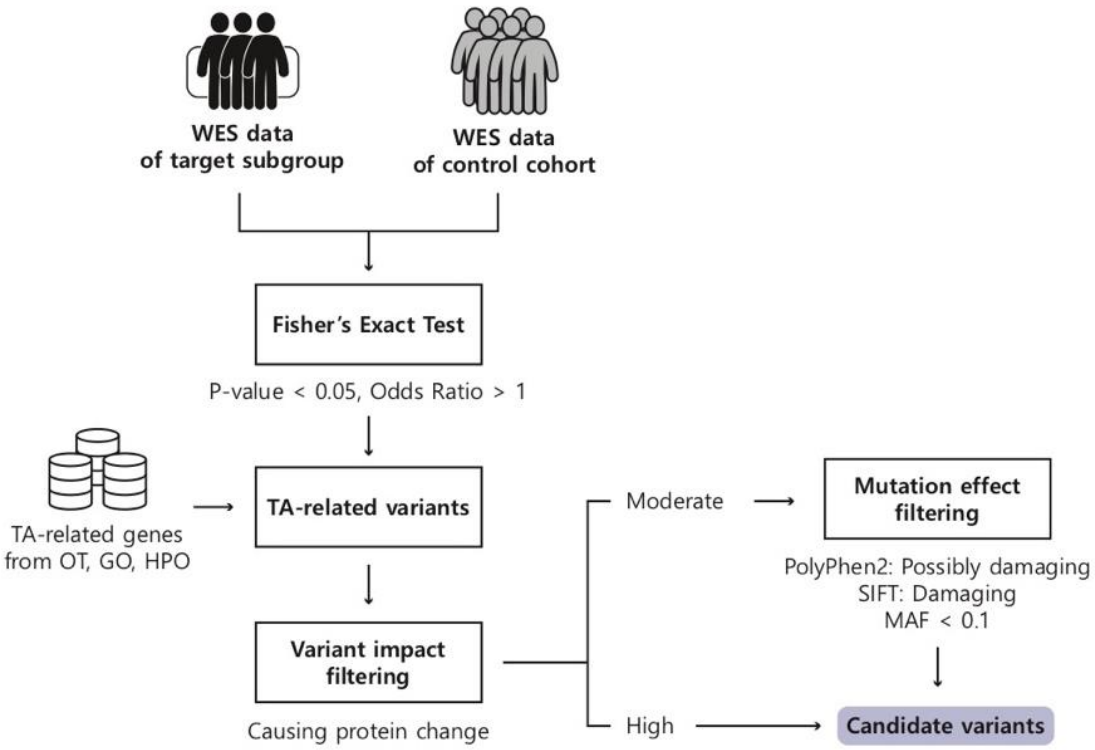


Figure 1. Flow chart detailing the variant filtration pipeline.

8. In silico mutation analysis

Sorting Intolerant From Tolerant (SIFT) [47] and Polymorphism Phenotyping v2 (PolyPhen2) [46] to detect mutations that may be responsible for TA. All possible non-redundant protein sequences in Ensembl database were analyzed for the two missense variants. SIFT determines the sequence conservation across multiple species, assuming mutations in highly conserved regions to be intolerable. PolyPhen2 predicts mutational impact by utilizing sequence co-evolution and protein structure. Table 6 summarizes the candidate genes with identified variants and their mutation impact analysis.

9. Gene set enrichment analysis

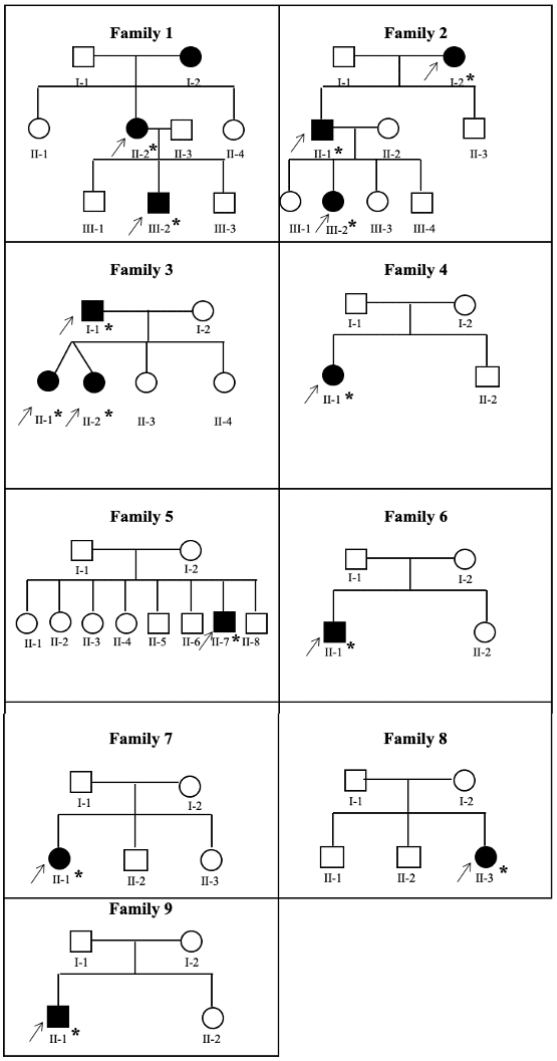
The g: Profiler [49] was used to perform gene set enrichment analysis (GSEA) for the candidate genes. The functional terms were derived from various data sources, including Gene Ontology [50] (GO molecular function, GO cellular component, and GO biological process), as well as biological pathways such as KEGG [51] and Reactome [52]). To increase the functional association of each gene, the gene sets were expanded by adding neighbor genes in the protein-protein interaction (PPI) network. The STRING (v11.5) [53] database was used for this process, with a confidence level of 700. The resulting network showed the relationship among the expanded gene sets, while the plot displays the function terms significantly enriched in the candidate genes at $p_{adj} < 0.01$.

III. RESULTS

1. Study subjects

The present study included nine Mongolian families, three of which were inherited families, as shown in Figure 2 a. WES analysis was conducted on a total of 14 samples, consisting of eight females and six males, with ages ranging from 9 to 69 years (Figure 2 b and Table 3). All 14 individuals were clinically diagnosed with non-syndromic TA, which was further confirmed through intra-oral examination and panoramic radiograph.

A



B

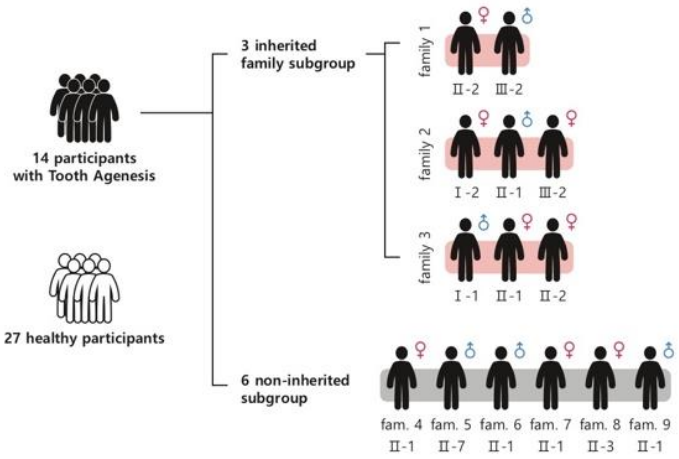


Figure 2. Pedigrees of the nine families and subgroup assignment of family members in the study. (**A**) Squares indicate males, and circles indicate females. Filled symbols represent individuals diagnosed with non-syndromic TA, while empty circles indicate unaffected subjects. Arrows indicate probands of each family. The asterisks represent subjects who underwent WES analysis. (**B**) Classification assignment of family members in subgroups.

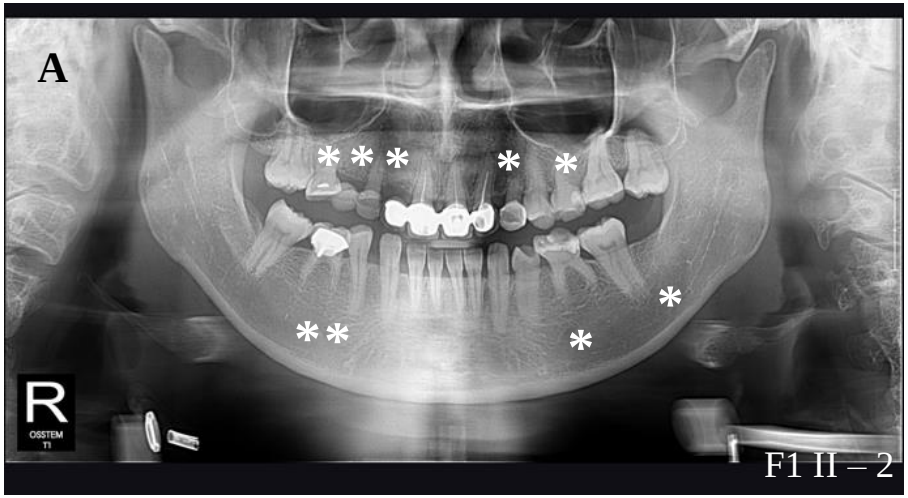
Table 3. The age and gender of participants.

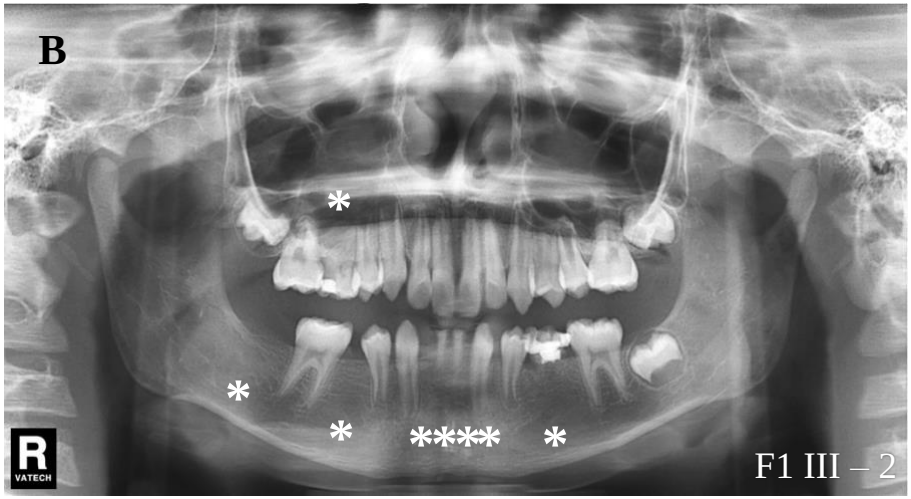
Family	Participant number	Age	Sex
1	II-2	30	F
1	III-2	12	M
2	I-2	69	F
2	II-1	42	M
2	III-2	15	F
3	I-1	46	M
3	II-1	15	F
3	II-2	15	F
4	II-1	9	F
5	II-7	12	M
6	II-1	13	M
7	II-1	19	F
8	II-3	9	F
9	II-1	14	M

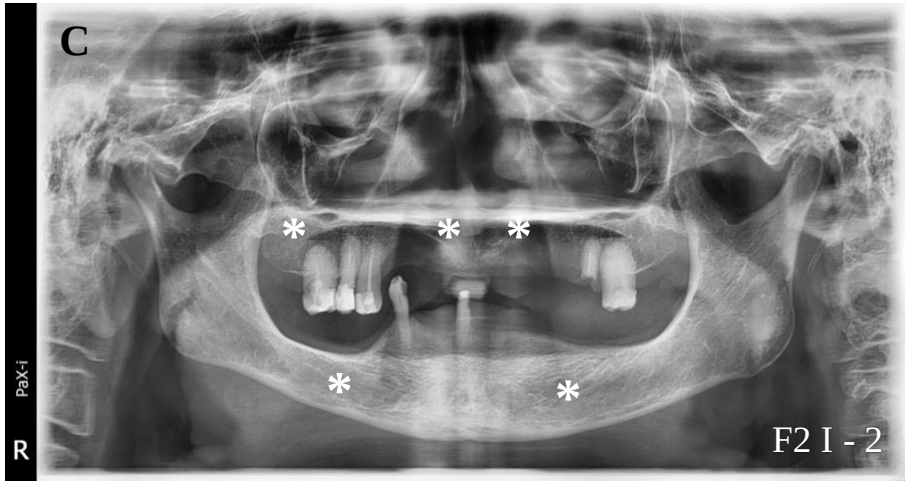
2. Clinical findings of families with non-syndromic TA

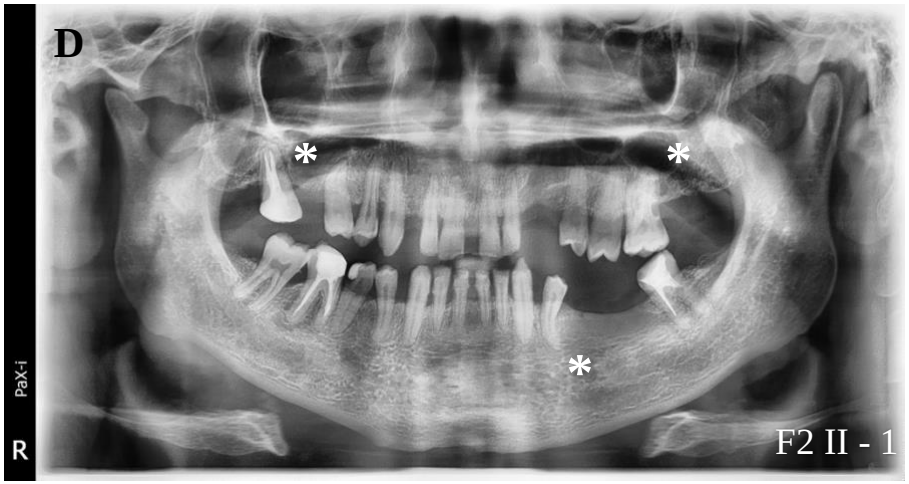
Family 1 consisted of a 30-year-old mother and a 12-year-old son as participants, with the mother (F1 II-2) missing nine permanent teeth (Figure 3 A) based on clinical and panoramic radiographs. The middle son (F1 III-2) was also found to be missing eight permanent teeth (Figure 3 B). Family 2 included a 69-year-old grandmother, a 42-year-old son, and a 15-year-old granddaughter as participants, where the grandmother (F2 I-2) had a total of five missing permanent teeth (Figure 3 C), her son (F2 II-1) was missing three permanent teeth (Figure 3 D), and her granddaughter (F2 III-2) was missing a significant number of teeth, specifically twenty-six permanent teeth (Figure 3 E). In Family 3, the participants were a 46-year-old father and his 15-year-old dizygotic twin daughters. The father (F3 I-1) was missing two permanent teeth (Figure 3 F), and one twin daughter (F3 II-1) missing six permanent teeth (Figure 3 G) while the other twin (F3 II-2) was missing four permanent teeth (Figure 3 H).

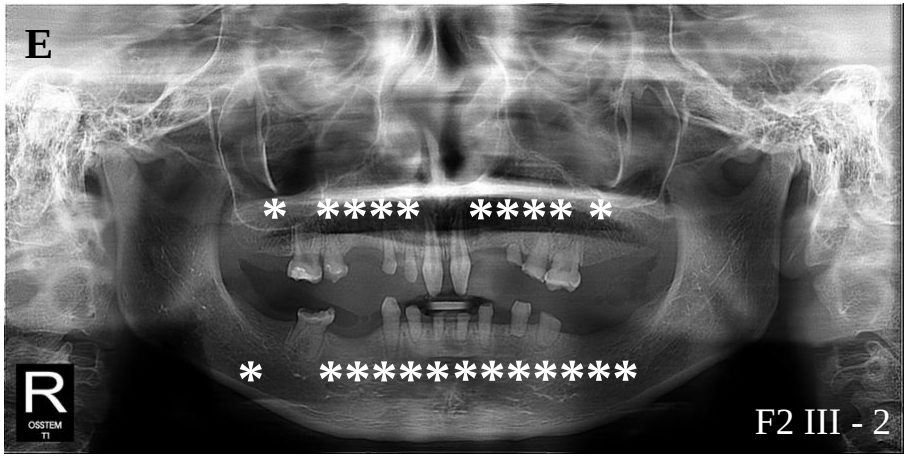
The remaining six families did not exhibit inherited patterns, and none of the family members of the participants showed any signs of TA. Family 4 included an 8-year-old girl (F4 II-1) was found to be missing seven permanent teeth (Figure 3 I), while in Family 5, an 11-year-old male (F5 II-7) was missing six permanent teeth (Figure 3 J). Family 6 had a 12-year-old boy (F6 II-1) who was missing twelve permanent teeth (Figure 3 K) and in Family 7, an 18-year-old girl (F7 II-1) was missing seven permanent teeth (Figure 3 L). Family 8 had an 8-year-old girl (F8 II-3) missing six permanent teeth (Figure 3 M); in Family 9, a 13-year-old boy (F9 II-1) was diagnosed with congenital agenesis and was missing a total of eleven permanent teeth (Figure 3 N). The participants who exhibit missing teeth, as indicated in the chart, are listed in Table 4.

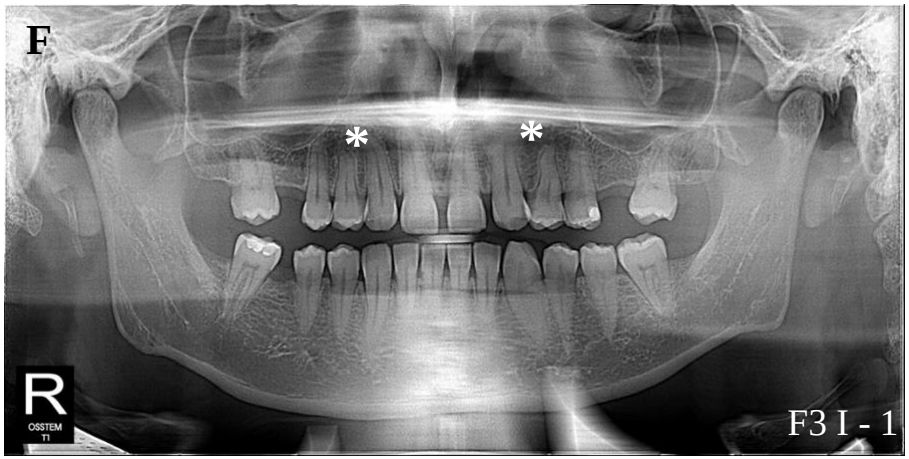


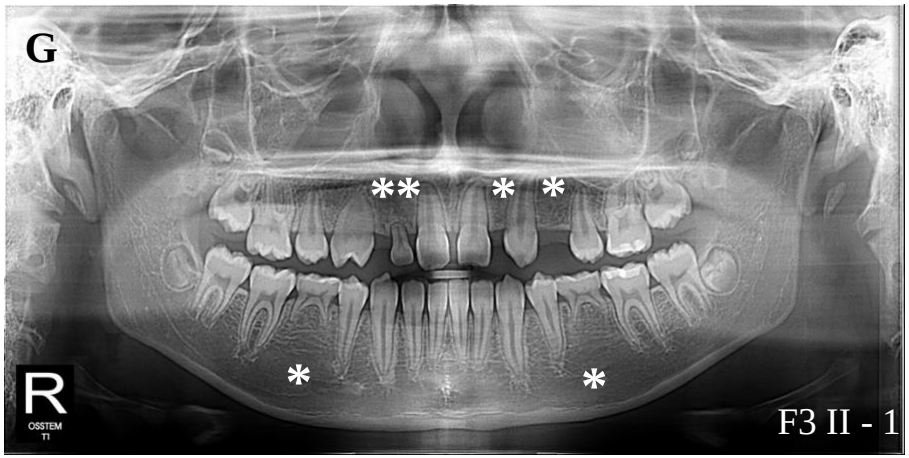


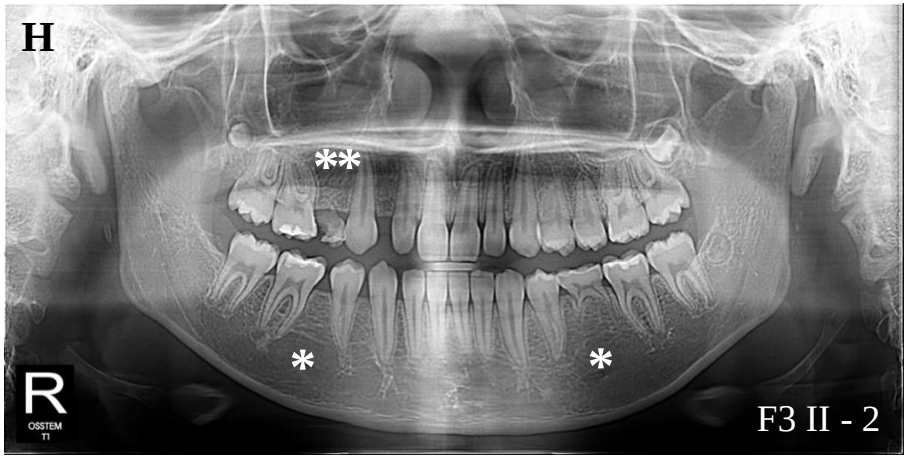


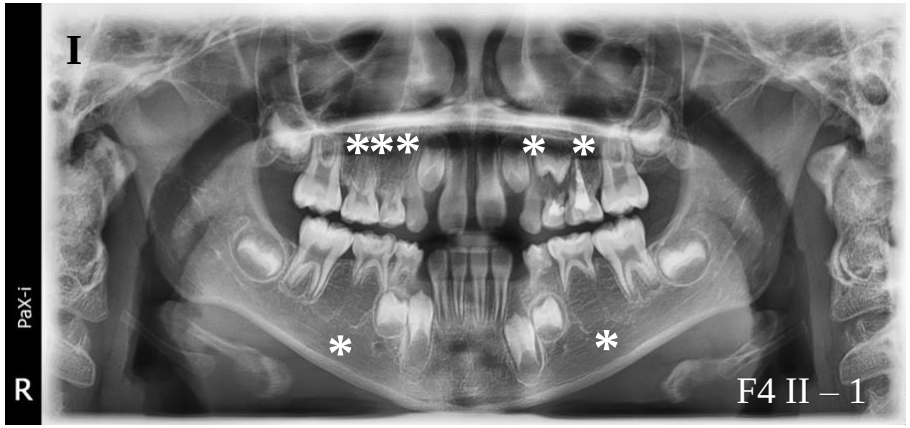


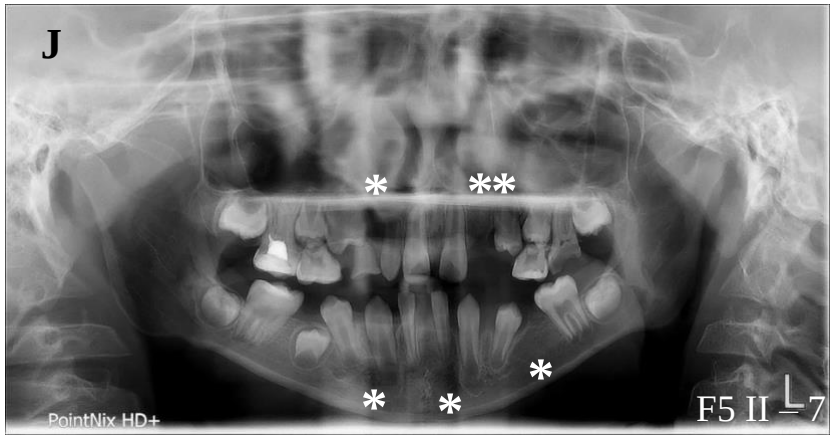


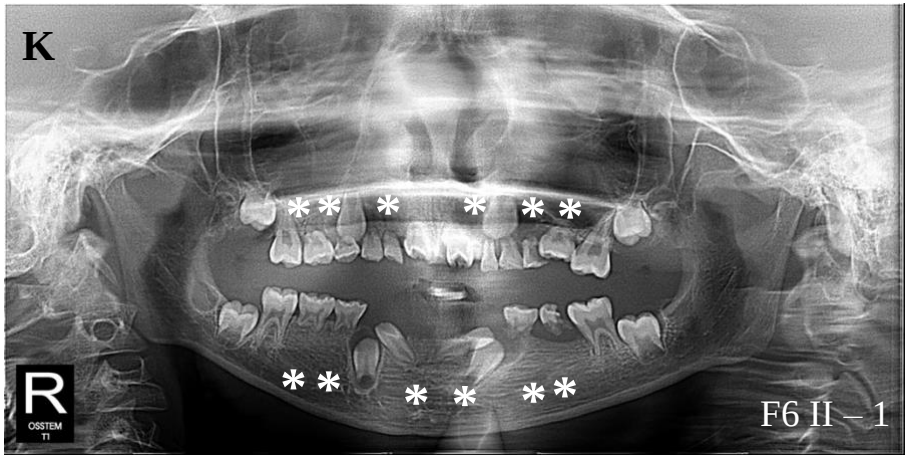


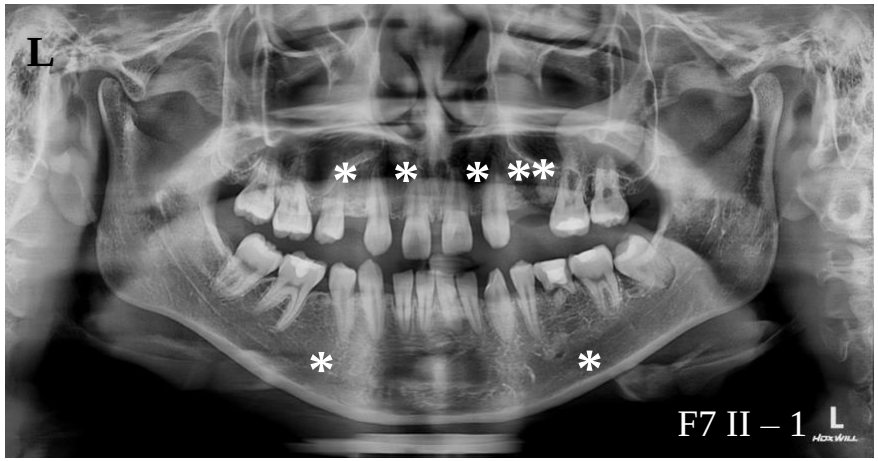












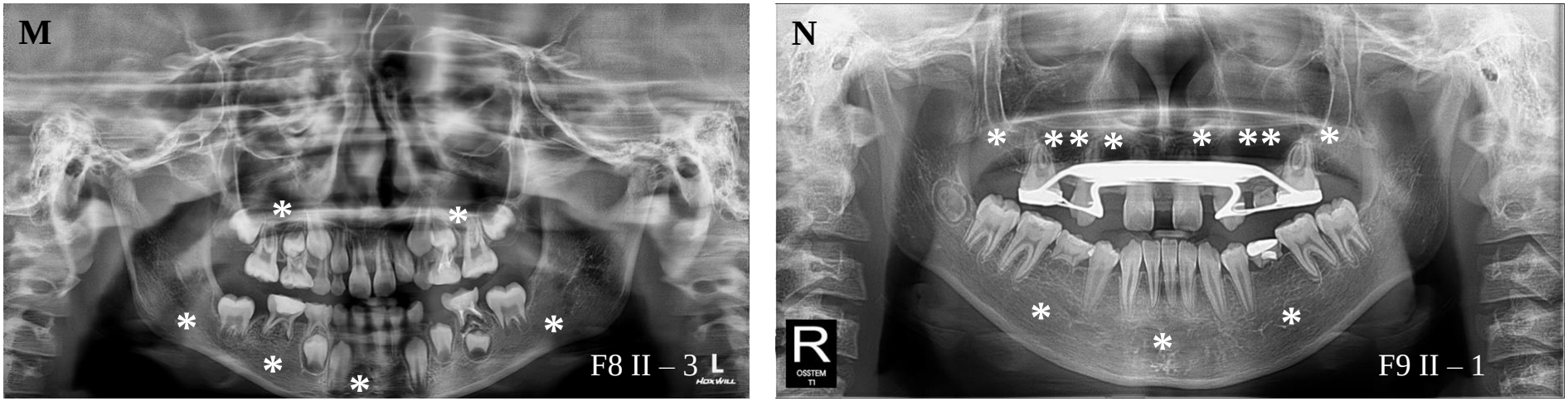


Figure 3. (**A** to **N**) Panoramic radiographs of 14 participants organized by their respective families illustrate the clinical findings of TA.

3. Candidate variants identified in WES analysis

WES was performed on all 14 participants who provided their consent. Total of five variants that met all the filtering criteria were identified through segregation analysis. Participants from inherited families with non-syndromic TA demonstrated autosomal dominant inheritance patterns, while those non-inherited families exhibited autosomal recessive inheritance (Table 4, 5). Among the Asian population, two rare missense variants were found: rs3850625 in CASNA1S with an Asian MAF of 0.0419 and rs12274923 in CDON with an Asian MAF of 0.0629. Additionally, a disruptive in-frame insertion in MAST4 (rs201910335) with an Asian MAF of 0.014 was also identified (Table 6).

Table 4. Phenotype (missing teeth) and genotypes of 5 variants from segregation analysis in 14 participants; minor risk allele in segregation analysis.

Family	Participants	Sex	TA Phenotype	Number of missing teeth	MAST4-C	ITGA-G	PITX2-G	CACNA1S-G	CDON-G
1	II-2	Female	Affected	9	**C/CGCT**	**G/GA**	**G/A**	G/G	**G/A**
1	III-2	Male	Affected	8	**C/CGCT**	**G/GA**	**G/A**	G/G	G/G
2	I-2	Female	Affected	5	C/C	G/G	**G/A**	G/G	G/G
2	II-1	Male	Affected	3	**C/CGCT**	G/G	**G/A**	G/G	**G/A**
2	III-2	Female	Affected	26	**C/CGCT**	G/G	**G/A**	G/G	**G/A**
3	I-1	Male	Affected	2	C/C	G/G	G/G	**G/A**	G/G
3	II-1	Female	Affected	6	**C/CGCT**	G/G	**G/A**	**G/A**	G/G
3	II-2	Female	Affected	4	**C/CGCT**	G/G	**G/A**	**G/A**	G/G
4	II-1	Female	Affected	7	C/C	G/G	**A/A**	G/G	**G/A**
5	II-7	Male	Affected	6	C/C	G/G	G/G	G/G	G/G
6	II-1	Male	Affected	12	C/C	G/G	G/G	G/G	G/G
7	II-1	Female	Affected	7	C/C	G/G	**G/A**	G/G	G/G
8	II-3	Female	Affected	6	C/C	G/G	**A/A**	G/G	**A/A**
9	II-1	Male	Affected	11	C/C	G/G	G/G	G/G	**G/A**

Table 5. A record of filtered variants. Gene, gene name; ID, variant id in SNP; CHR, chromosome; POS, base pair position; REF, reference allele; ALT, alternative allele.

Gene	ID	CHR	POS	REF	ALT	Transcript ID	Protein change	Impact
MAST4	rs201910335	5	65892764	C	CGCT	ENST00000404260	Leu95dup	MODERATE
ITGA6	-	2	173337540	G	GA	ENST00000409532	Asp114fs	HIGH
PITX2	rs2278782	4	111542154	G	A	ENST00000557119	Gln193*	HIGH
CACNA1S	rs3850625	1	201016296	G	A	ENST00000367338	Arg1520Cys	MODERATE
CDON	rs12274923	11	125871715	G	A	ENST00000531738	Ala63Val	MODERATE

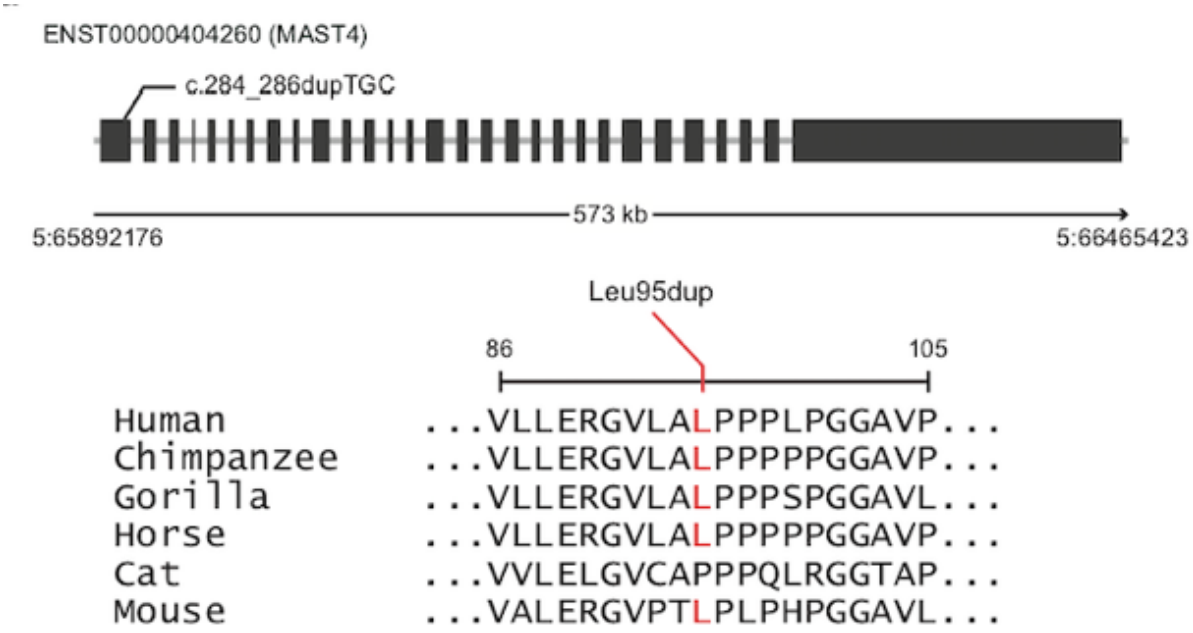
4. In silico mutation impact analysis of the variants

Two different computational tools were applied to analyze the impact of mutations in MAST4, CACNA1S, and CDON with filtered variants: SIFT [54] and PolyPhen2 [55]. For the two missense variants, canonical Ensembl sequences were analyzed (Table 6). The SIFT predicted a damaging effect on a missense variant in CACNA1S, while PolyPhen2 did not. In contrast, a variant in CDON was predicted to be possibly damaging by PolyPhen2 and had both damaging and tolerant prediction by SIFT. The variants were found to be located in evolutionarily conserved positions and might be under intense selective pressure. For instance, a missense variant in CACNA1S (located in exon 7, and arginine at the amino acid position 1520) and a variant in CDON (located in exon 10, and alanine at the amino acid position 63) were well-conserved in mammal orthologs (Figure 4 B, C). In the case of MAST4, the mutation has occurred at the site where sequences are well-conserved across mammal orthologs (Figure 4 A). To assess the gene tolerance to mutations concerning TA, Loss-of-Function observed/expected upper bound fraction (LOEUF) scores were measured. A lower LOEUF score indicates that a gene is more susceptible to damage by mutations and may have a more severe impact on affected individuals [56]. The findings revealed that MAST4 had a relatively low LOEUF score of 0.38 and that other genes also had LOEUF scores below 1(CACNA1S had a LOEUF score of 0.56 and CDON had a LOEUF score of 0.96), suggesting that the gene functions might be damaged by the variants (Table 6).

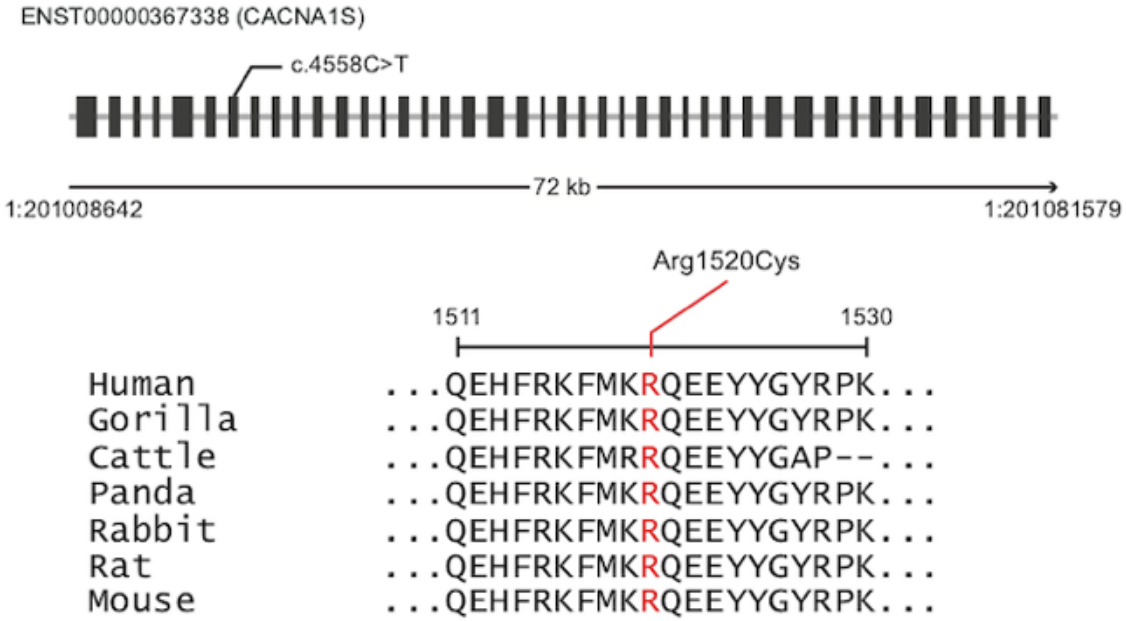
Table 6. Analysis of mutation impact on candidate genes; results of PolyPhen2 (B: benign, P: possibly damage) and SIFT (D: damage, T: tolerate). Effect; defining mutations; LOEUF, indicates the inverse score of loss-of-function intolerance; ASN MAF, alternative allele frequency in 1000Gp1 Asian descendent samples; MAF, alternative allele frequency in whole 1000Gp1 data.

Gene	UniProt ID	Effect	PolyPhen2	SIFT	LOEUF	MAF	ASN MAF
MAST4	O15021	disruptive_inframe_insertion	-	-	0.38	0.005814	0.01434
CACNA1S	Q13698	missense_variant	B	D, D	0.56	0.070971	0.041958
CDON	Q4KMG0	missense_variant	P, P, P	T, D, T	0.96	0.125916	0.062937

A



B



C

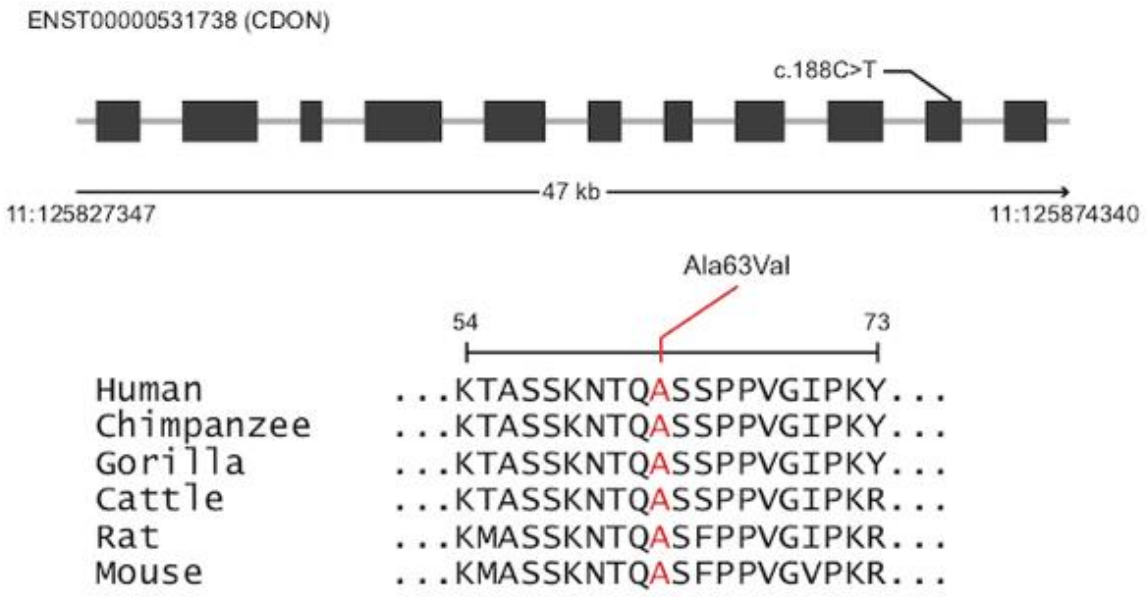
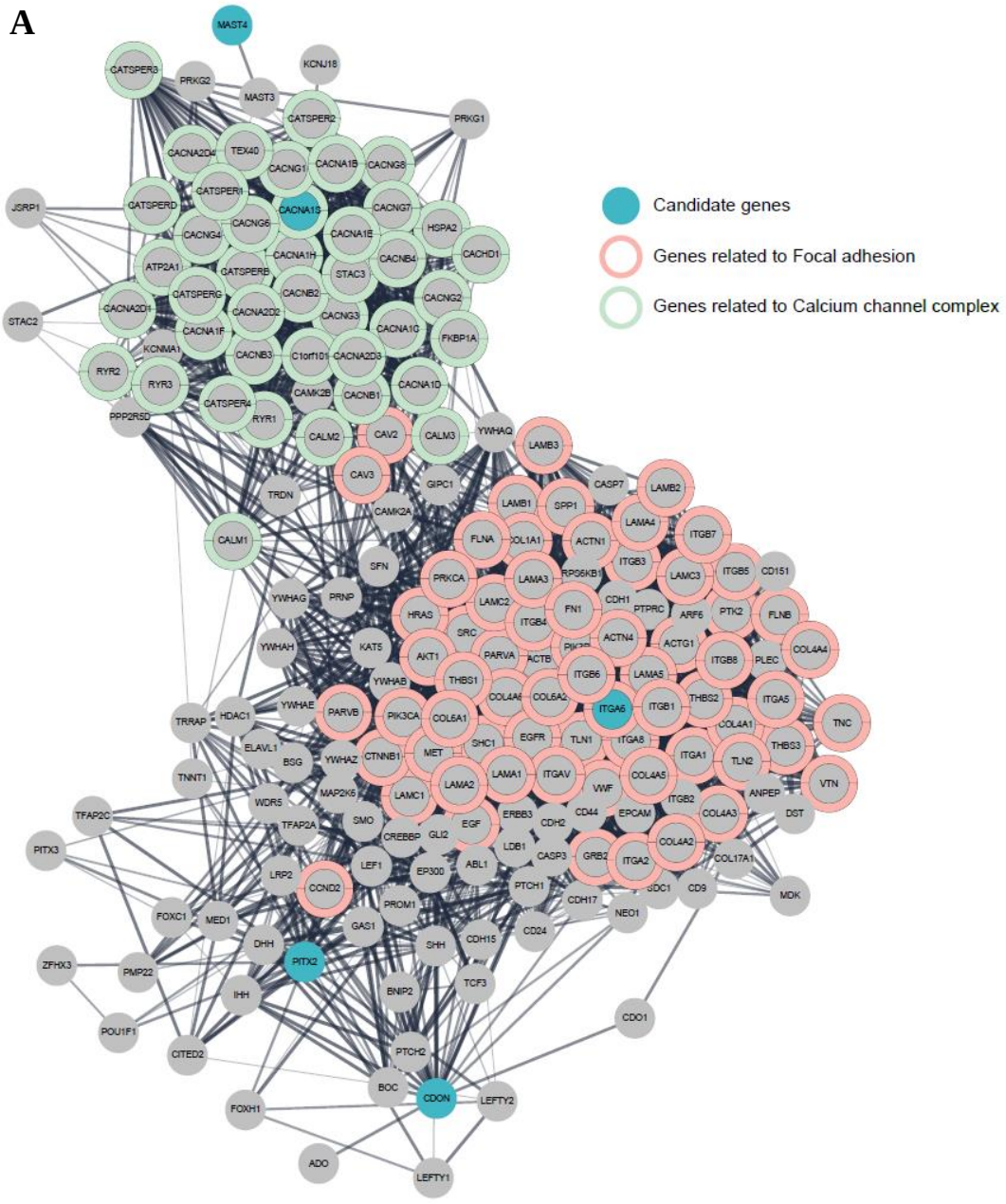


Figure 4. Mutation impact analysis of MAST4, CACNA1S, and CDON. (**A**) The amino acid sequence alignment of MAST4 in various mammalian orthologs revealed a mutated residue of Leu95dup, which is highlighted in red. (**B**) The amino acid sequence alignment of CACNA1S in various mammalian orthologs revealed a mutated residue of Arg1520Cys, which is highlighted in red. (**C**) The amino acid sequence alignment of CDON in various mammalian orthologs revealed a mutated residue of Ala63Val, which is highlighted in red.

5. Gene set enrichment analysis

Next, to comprehend the functional contribution of candidate genes in non-syndromic TA, an investigation of gene set enrichment analysis (GSEA) was performed (Table 7). This analysis uncovered three noteworthy categories, namely Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome pathway (REAC). For the GSEA, the list of genes with neighbors in the PPI network was expanded. Using the five candidate variant genes, function terms related to focal adhesion and calcium channel complex were highly ranked and formed functional gene clusters in the PPI network (Figure 5).



B

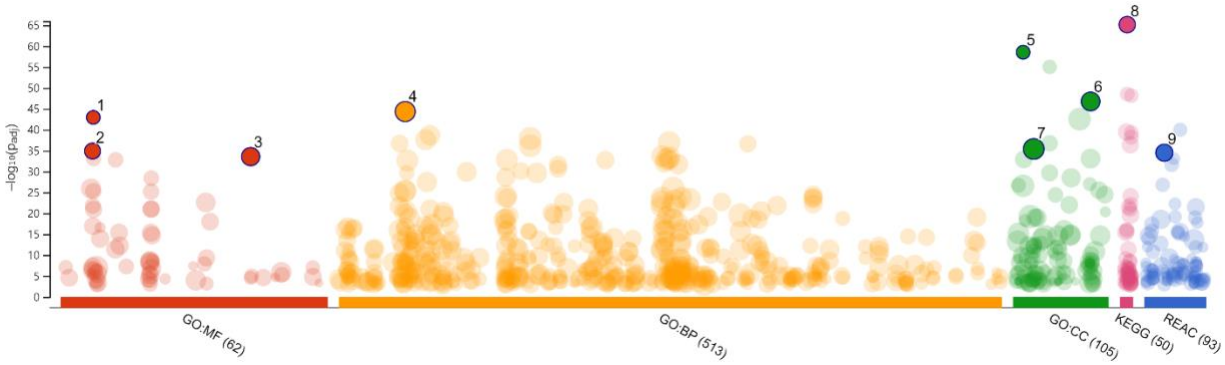


Figure 5. Gene set enrichment analysis result of candidate genes and its PPI network neighboring genes. (**A**) The PPI network of expanded gene set. Cyan-colored nodes are the five candidate genes. (**B**) Adjusted p-values of function terms as the result of GSEA. Highlighted dots are the high-ranked function terms related to the focal adhesion and calcium channel complex.

Table 7. List of highlighted function terms of the gene set enrichment analysis result plot.

	Source	Term	Term ID	Count	p-value
1	GO:MF	voltage-gated calcium channel activity	GO:0005245	28	1.10E-43
2	GO:MF	integrin binding	GO:0005178	35	1.23E-35
3	GO:MF	cell adhesion molecule binding	GO:0050839	52	2.66E-34
4	GO:BP	cell adhesion	GO:0007155	87	4.42E-45
5	GO:CC	voltage-gated calcium channel complex	GO:0005891	34	2.80E-59
6	GO:CC	plasma membrane protein complex	GO:0098797	66	1.69E-47
7	GO:CC	cell junction	GO:0030054	87	4.01E-36
8	KEGG	Focal adhesion	KEGG:04510	67	6.90E-66
9	REAC	Extracellular matrix organization	REAC: R-HSA-1474244	50	3.11E-35

IV. DISCUSSION

To the best of current knowledge, this is the first study to use a combination of WES and bioinformatic analysis to comprehensively investigate the genetic factors underlying functional mechanisms associated with TA in Mongolian families.

TA is a congenital disorder characterized by the absence of one or more teeth due to the failure to develop. The underlying mechanism of TA is complex and still needs to be fully understood [19]. In recent years, there has been a growing interest in utilizing WES and bioinformatic analysis to identify pathogenic variants in candidate genes associated with TA [36]. WES is a highly effective method for identifying genetic variants associated with genetic diseases, and genes involved in the developmental and differentiation of tooth germ cells have been discovered using this method [31].

Although new variants are still being discovered, the exact pathogenic mechanism underlying TA remains unclear. Previous studies on TA have primarily focused on genetic variants, with a limited investigation into the functional aspects such as protein function or molecular pathways. This study aimed to comprehend TA by investigating not only the genetic variants but also the underlying functional mechanisms involved in the process. To contribute to this understanding, a bioinformatic analysis was conducted. This encompassed comparing identified genetic variants with public databases and literature, as well as an assessment of the frequency of the variant in the general population (LOEUF analysis), a lower minor allele frequency in the Asian population (ASN MAF) and a prediction of the potential impact of the variant on protein function (SIFT and PolyPhen analyses).

In this study, impact variants in MAST4 were identified. A disruptive in-frame insertion in MAST4 with an Asian MAF of 0.014 indicates that this particular variant is rare with Asian population. The rarity of this insertion in the population underscores its potential significance, warranting further investigation into its functional implications and potential associations with

relevant phenotypic traits or diseases. Understanding the prevalence and impact of such variants contributes valuable insights to knowledge of genetic diversity and susceptibility within distinct populations. The result of variant impact analysis in the general population revealed that MAST4 showed a lower LOEUF score of 0.38 (Table 6) compared to CACNA1S and CDON, suggesting it could be related to the tooth damage and have a more severe impact on affected individuals.

Indeed, a disruptive variant (rs201910335) in the MAST4 gene was identified in 6 individuals from three inherited families with TA (Table 4). The MAST4 gene was first characterized in 2006 through bioinformatics analysis [57]. Studies have investigated its roles in crucial biological processes, including cell cycle progression, cell migration, and neuronal development [58]. The precise mechanism by which the MAST4 gene leads to TA is not fully understood up to this point. However, a recent study suggests that MAST4 is closely involved in amelogenesis process of mouse incisors and may serve as a critical regulator of Amelogenesis Imperfecta [59] by deactivating Wnt/β-catenin signaling pathway [60].

This result is particularly noteworthy and captures researchers attention because the Wnt signaling pathway is also a major pathway in TA [61, 33]. During tooth formation, Wnt signaling pathway is activated in the dental epithelium, which leads to the formation of a structure called the enamel knot. The enamel knot serves as a signaling center that helps direct the dental mesenchyme's growth and differentiation [33]. Mutations or alterations in the genes involved in Wnt signaling pathway, such as WNT10A, AXIN2, and LRP6, are known to be associated with TA. Mutations in WNT10A gene, which encodes a ligand for the canonical Wnt pathway, have been linked to TA [32]. The mutations in AXIN2, which is negative regulator of the gene canonical Wnt pathway, have been also been associated with TA, as well as other dental abnormalities [62]. In addition, variations in the LRP6 gene, which encodes a co-receptor for the canonical Wnt pathway, have been linked to TA [63].

These findings highlight the significance of Wnt and Wnt-associated signaling pathways in the genetic etiology of TA and may facilitate the identification of novel genes associated with this condition. The results suggest that the MAST4 gene's relationship with the Wnt pathway may significantly contribute to the development of TA. However, further research is necessary to comprehensively elucidate the underlying molecular mechanism through which the MAST4 and related pathways influence tooth development and contribute to the pathogenesis of TA.

A stop-gained mutation in PITX2 gene (rs2278782) was identified in 8 participants from both inherited and non-inherited families (Table 4). This finding is consistent with previous research linking PITX2 mutation to non-syndromic TA and dental anomalies [39, 64]. PITX2 is a transcription factor for proper tooth development by activating target genes through the Wnt signaling pathway, involving β-catenin [65]. Mutation on the PITX2 gene can disrupt the proper secretion of Wnt4, Wnt6, and Wnt10 from dental epithelium, resulting in a dysfunctional enamel knot leading to arrest of tooth development due to the absence of Wnt/β-catenin activity [66-68].

Variants in CDON and ITGA6 genes were also found, which are co-receptors for the SHH signaling pathway crucial for tooth development [69, 70]. The SHH signaling affects cell polarization in the early tooth bud, determining the number of teeth in the permanent dentition. It also interacts with other pathways, including Wnt, to ensure proper tooth development [71]. Therefore, these findings indicate a connection between the Wnt signaling pathway and TA.

Gene set enrichment analysis revealed the functional terms of five candidate genes associated with focal adhesion and calcium channel complex to be highly important, with two gene clusters linked by the candidate genes and their neighboring genes in the PPI network. Focal adhesions are specialized structures that allow cells to interact with extracellular matrix (ECM) through multiprotein complexes, mediating cell adhesion, migration, and signaling processes [72, 73] necessary for the formation of the tooth germ, enamel, dentin, and periodontal ligament [74-76]. Inhibiting focal adhesion kinase (FAK) can hinder the bud-to-

cap morphogenesis process, potentially leading to TA [77, 78]. Calcium signaling, regulated by voltage-gated channels (VDCC) [79] and important in tooth development [80], can be disrupted by mutations in calcium channel genes, such as CACNA1S, leading to dental malformations and abnormalities [81]. These findings indicate that disruptions in focal adhesion may affect tooth morphogenesis and differentiation, which may lead to TA. Furthermore, the significance of calcium signaling in tooth development suggests that calcium channelopathies can contribute to dental abnormalities.

The selection of a Korean control group from Ansan-Ansung population in this study is rooted in several key considerations. Firstly, Mongolia's current lacks of a comprehensive genetic databank presents a challenge in conducting comparative genetic studies aimed at identifying variants associated with non-syndromic TA within the Mongolian population. Secondly, despite Mongolia's absence of a dedicated databank, previous research has highlighted shared genetic similarities among East Asian populations, including both Koreans and Mongolians. Studies have consistently indicated commonalities in genetic inheritance patterns across ethnic groups in East Asia [82]. However, further studies are needed to explore specific genetic variations within the Mongolian population itself.

In light of the limitation of this study, specifically the inclusion of participants with varying degrees of TA in the same analysis, it would be beneficial for future investigations to stratify participants into distinct subgroups based on the severity of TA. This approach would allow for a more refined analysis and deeper understanding of the specific mutation types and genetic factors associated with each form of TA. Furthermore, WES did not consider the unaffected family members due to specific reasons such as logistical constraints, ethical considerations, and financial limitations, which led to their inability to participate in this study. Additionally, the relatively small sample size in this study may have limited the statistical power to detect smaller effects or associations accurately. While researchers have made efforts to address these limitations through bioinformatic analysis and functional analysis, larger and more

homogeneous cohorts are needed to further validate and expand upon these findings. Nonetheless, despite these limitations, this study provides valuable insights into the genetic basis of TA and identifies potential candidate genes associated with this condition.

The study has identified novel candidate genes and variants that contribute to the occurrence of non-syndromic TA in Mongolian families. Furthermore, the results emphasize the significance of Wnt signaling pathway genes, as well as focal adhesion and calcium channel complexes, in regulating the process of tooth development. The identification of novel pathogenic genes associated with TA can enhance the understanding of the disease's molecular mechanisms. This could lead to improved diagnosis, prevention, and treatment strategies. Moreover, these findings may guide clinical practice, inform clinical gene diagnosis, and facilitate the development of targeted treatment options for individuals affected by TA. In clinical practice, early detection of TA through biomarkers significantly enhances dental management and facilitates orthodontic and prosthetic treatments.

V. CONCLUSION

In conclusion, novel candidate genes and genetic variants contributing to non-syndromic TA in Mongolian families have been identified. The findings emphasize the importance of the Wnt signaling pathway and its associations in the molecular development of non-syndromic TA. Additionally, the research provides evidence for the involvement of focal adhesion and calcium channel complexes in the regulation of the tooth development process.

VI. REFERENCES

1. Du, R., N. Dinckan, X. Song, et al. (2018) Identification of likely pathogenic and known variants in TSPEAR, LAMB3, BCOR, and WNT10A in four Turkish families with tooth agenesis. Hum Genet 137:689-703.
2. Zhou, M., H. Zhang, H. Camhi, et al. (2021) Analyses of oligodontia phenotypes and genetic etiologies. Int J Oral Sci 13:32.
3. Letra, A. (2022) Rethinking the Genetic Etiology of Nonsyndromic Tooth Agenesis. Curr Osteoporos Rep 20:389-397.
4. Park, H., J.S. Song, T.J. Shin, et al. (2019) WNT10A mutations causing oligodontia. Arch Oral Biol 103:8-11.
5. Bailleul-Forestier, I., M. Molla, A. Verloes, et al. (2008) The genetic basis of inherited anomalies of the teeth. Part 1: clinical and molecular aspects of non-syndromic dental disorders. Eur J Med Genet 51:273-91.
6. Huang, X.F. and Y. Chai (2012) Molecular regulatory mechanism of tooth root development. Int J Oral Sci 4:177-81.
7. Jo, C.W., J.H. Lee, J.S. Song, et al. (2021) Isolated and Sporadic Human Mesiodens Is Associated with a Synonymous Variant in the ACVR2A Gene. Pediatr Dent 43:39-43
8. Bonczek, O., P. Bielik, P. Krejci, et al. (2018) Next generation sequencing reveals a novel nonsense mutation in MSX1 gene related to oligodontia. PLoS One 13:e0202989.
9. Matalova, E., J. Fleischmannova, P.T. Sharpe, et al. (2008) Tooth agenesis: from molecular genetics to molecular dentistry. J Dent Res 87:617-23. 10.1177/154405910808700715
10. Carter, K. and S. Worthington (2015) Morphologic and Demographic Predictors of Third Molar Agenesis: A Systematic Review and Meta-analysis. J Dent Res 94:886-94. 10.1177/0022034515581644
11. Polder, B.J., M.A. Van't Hof, F.P. Van der Linden, et al. (2004) A meta-analysis of the prevalence of dental agenesis of permanent teeth. Community Dent Oral Epidemiol 32:217-26. 10.1111/j.1600-0528.2004.00158.x

12. Howe, B.J., M.E. Cooper, A.R. Vieira, et al. (2015) Spectrum of Dental Phenotypes in Nonsyndromic Orofacial Clefting. J Dent Res 94:905-12. 10.1177/0022034515588281
13. Mues, G.I., R. Griggs, A.J. Hartung, et al. (2009) From ectodermal dysplasia to selective tooth agenesis. Am J Med Genet A 149A:2037-41.
14. Bock, N.C., S. Lenz, G. Ruiz-Heiland, et al. (2017) Nonsyndromic oligodontia : Does the Tooth Agenesis Code (TAC) enable prediction of the causative mutation? J Orofac Orthop 78:112-120.
15. Bonczek, O., P. Krejci, L. Izakovicova-Holla, et al. (2021) Tooth agenesis: What do we know and is there a connection to cancer? Clin Genet 99:493-502.
16. Ye, X. and A.B. Attaie (2016) Genetic Basis of Nonsyndromic and Syndromic Tooth Agenesis. J Pediatr Genet 5:198-208. 10.1055/s-0036-1592421.
17. Al-Ani, A.H., J.S. Antoun, W.M. Thomson, et al. (2017) Hypodontia: An Update on Its Etiology, Classification, and Clinical Management. Biomed Res Int 2017:9378325.
18. Lan, R., Y. Wu, Q. Dai, et al. (2022) Gene mutations and chromosomal abnormalities in syndromes with tooth agenesis. Oral Dis.
19. Sun, K., M. Yu, I. Yeh, et al. (2021) Functional study of novel PAX9 variants: The paired domain and non-syndromic oligodontia. Oral Dis 27:1468-1477.
20. Yin, W. and Z. Bian (2015) The Gene Network Underlying Hypodontia. J Dent Res 94:878-85. 10.1177/0022034515583999
21. 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. 10.1177/0022034518777460
22. Keskin, G., K. Karaer, and Z. Ucar Gundogar (2022) Targeted next-generation sequencing (NGS) analysis of mutations in nonsyndromic tooth agenesis candidate genes : Analysis of a Turkish cohort. J Orofac Orthop 83:65-74. 10.1007/s00056-021-00284-4
23. Creton, M., M.J. van den Boogaard, T. Maal, et al. (2013) Three-dimensional analysis of tooth dimensions in the MSX1-missense mutation. Clin Oral Investig 17:1437-45.
24. Yu, M., H. Wang, Z. Fan, et al. (2019) BMP4 mutations in tooth agenesis and low bone mass. Arch Oral Biol 103:40-46.

25. Wong, S., H. Liu, B. Bai, et al. (2014) Novel missense mutations in the AXIN2 gene associated with non-syndromic oligodontia. Arch Oral Biol 59:349-53.
26. Nikopensius, T., T. Annilo, T. Jagomagi, et al. (2013) Non-syndromic tooth agenesis associated with a nonsense mutation in ectodysplasin-A (EDA). J Dent Res 92:507-11. 10.1177/0022034513487210
27. Zhang, H., X. Kong, J. Ren, et al. (2021) A novel EDAR missense mutation identified by whole-exome sequencing with non-syndromic tooth agenesis in a Chinese family. Mol Genet Genomic Med 9:e1684.
28. Martinez-Romero, M.C., M.J. Ballesta-Martinez, V. Lopez-Gonzalez, et al. (2019) EDA, EDAR, EDARADD and WNT10A allelic variants in patients with ectodermal derivative impairment in the Spanish population. Orphanet J Rare Dis 14:281.
29. Yu, P., W. Yang, D. Han, et al. (2016) Mutations in WNT10B Are Identified in Individuals with Oligodontia. Am J Hum Genet 99:195-201.
30. Wang, H., Y. Liu, Y. Zheng, et al. (2021) A novel missense mutation of LRP6 identified by whole-exome sequencing in a Chinese family with non-syndromic tooth agenesis. Orthod Craniofac Res 24:233-240.
31. Yu, K., J. Dou, W. Huang, et al. (2022) Expanding the genetic spectrum of tooth agenesis using whole-exome sequencing. Clin Genet 102:503-516.
32. Dinckan, N., R. Du, L.E. Petty, et al. (2018) Whole-Exome Sequencing Identifies Novel Variants for Tooth Agenesis. J Dent Res 97:49-59.
33. Yu, M., S.W. Wong, D. Han, et al. (2019) Genetic analysis: Wnt and other pathways in nonsyndromic tooth agenesis. Oral Dis 25:646-651.
34. Yue, H., J. Liang, G. Song, et al. (2022) Mutation analysis in patients with nonsyndromic tooth agenesis using exome sequencing. Mol Genet Genomic Med 10:e2045. 10.1002/mgg3.2045
35. Yang, L., J. Liang, H. Yue, et al. (2020) Two novel mutations in MSX1 causing oligodontia. PLoS One 15:e0227287.
36. Salvi, A., E. Giacopuzzi, E. Bardellini, et al. (2016) Mutation analysis by direct and whole exome sequencing in familial and sporadic tooth agenesis. Int J Mol Med 38:1338-1348.
37. Tallón-Walton, V., M.C. Manzanares-Céspedes, P. Carvalho-Lobato, et al. (2014) Exclusion of PAX9 and MSX1 mutation in six families affected by

tooth agenesis. A genetic study and literature review. Med Oral Patol Oral Cir Bucal 19:e248-54. 10.4317/medoral.19173

38. Wong, S.W., H.C. Liu, D. Han, et al. (2014) A novel non-stop mutation in MSX1 causing autosomal dominant non-syndromic oligodontia. Mutagenesis 29:319-23. 10.1093/mutage/geu019
39. Fan, Z., S. Sun, H. Liu, et al. (2019) Novel PITX2 mutations identified in Axenfeld-Rieger syndrome and the pattern of PITX2-related tooth agenesis. Oral Dis 25:2010-2019.
40. Liu, Z., L. Zhu, R. Roberts, et al. (2019) Toward Clinical Implementation of Next-Generation Sequencing-Based Genetic Testing in Rare Diseases: Where Are We? Trends Genet 35:852-867.
41. Ulintz, P.J., W. Wu, and C.M. Gates (2019) Bioinformatics Analysis of Whole Exome Sequencing Data. Methods Mol Biol 1881:277-318. 10.1007/978-1-4939-8876-1_21
42. Ochoa, D., A. Hercules, M. Carmona, et al. (2023) The next-generation Open Targets Platform: reimagined, redesigned, rebuilt. Nucleic Acids Res 51:D1353-d1359. 10.1093/nar/gkac1046
43. (2021) The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res 49:D325-d334. 10.1093/nar/gkaa1113
44. Köhler, S., M. Gargano, N. Matentzoglu, et al. (2021) The Human Phenotype Ontology in 2021. Nucleic Acids Res 49:D1207-d1217. 10.1093/nar/gkaa1043
45. Cingolani, P., A. Platts, L. Wang le, et al. (2012) 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 (Austin) 6:80-92.
46. Flanagan, S.E., A.M. Patch, and S. Ellard (2010) Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. Genet Test Mol Biomarkers 14:533-7.
47. Ng, P.C. and S. Henikoff (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31:3812-4.
48. Linck, E. and C.J. Battey (2019) Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. Mol Ecol Resour 19:639-647.

49. Reimand, J., R. Isserlin, V. Voisin, et al. (2019) Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc 14:482-517.
50. Kabir, A. and A. Shehu (2022) GOProFormer: A Multi-Modal Transformer Method for Gene Ontology Protein Function Prediction. Biomolecules 12.
51. Chen, L., Y.H. Zhang, G. Lu, et al. (2017) Analysis of cancer-related lncRNAs using gene ontology and KEGG pathways. Artif Intell Med 76:27-36.
52. Mi, H., X. Huang, A. Muruganujan, et al. (2017) PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res 45:D183-d189.
53. Szklarczyk, D., A.L. Gable, D. Lyon, et al. (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47:D607-d613.
54. Kumar, P., S. Henikoff, and P.C. Ng (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4:1073-81.
55. Adzhubei, I., D.M. Jordan, and S.R. Sunyaev (2013) Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet Chapter 7:Unit7.20.
56. Seaby, E.G., N.S. Thomas, A. Webb, et al. (2023) Targeting de novo loss-of-function variants in constrained disease genes improves diagnostic rates in the 100,000 Genomes Project. Hum Genet 142:351-362.
57. Sun, L., S. Gu, X. Li, et al. (2006) [Identification of a novel human MAST4 gene, a new member of the microtubule associated serine-threonine kinase family]. Mol Biol (Mosk) 40:808-15.
58. Gongol, B., T.L. Marin, J.D. Jeppson, et al. (2017) Cellular hormetic response to 27-hydroxycholesterol promotes neuroprotection through AICD induction of MAST4 abundance and kinase activity. Sci Rep 7:13898.
59. Lee, D.-J., P. Kim, H.-Y. Kim, et al. (2021) Unraveling the roles of Mast4 in amelogenesis via regulating DLX3 and stem cell maintenance of mouse incisors. bioRxiv.

60. Zhan, Y., X. Li, X. Gou, et al. (2018) DLX3 Inhibits the Proliferation of Human Dental Pulp Cells Through Inactivation of Canonical Wnt/β-Catenin Signaling Pathway. Front Physiol 9:1637.
61. Lan, Y., S. Jia, and R. Jiang (2014) Molecular patterning of the mammalian dentition. Semin Cell Dev Biol 25-26:61-70.
62. 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 74:1043-50.
63. Ockeloen, C.W., K.D. Khandelwal, K. Dreesen, et al. (2016) Novel mutations in LRP6 highlight the role of WNT signaling in tooth agenesis. Genet Med 18:1158-1162.
64. Intarak, N., T. Theerapanon, C. Ittiwut, et al. (2018) A novel PITX2 mutation in non-syndromic orodental anomalies. Oral Dis 24:611-618.
65. Amen, M., X. Liu, U. Vadlamudi, et al. (2007) PITX2 and beta-catenin interactions regulate Lef-1 isoform expression. Mol Cell Biol 27:7560-73.
66. Lin, C.R., C. Kioussi, S. O'Connell, et al. (1999) Pitx2 regulates lung asymmetry, cardiac positioning and pituitary and tooth morphogenesis. Nature 401:279-82.
67. St Amand, T.R., Y. Zhang, E.V. Semina, et al. (2000) Antagonistic signals between BMP4 and FGF8 define the expression of Pitx1 and Pitx2 in mouse tooth-forming anlage. Dev Biol 217:323-32.
68. Zhu, X., P. Zhao, Y. Liu, et al. (2013) Intra-epithelial requirement of canonical Wnt signaling for tooth morphogenesis. J Biol Chem 288:12080-9.
69. Hermans, F., L. Hemeryck, I. Lambrichts, et al. (2021) Intertwined Signaling Pathways Governing Tooth Development: A Give-and-Take Between Canonical Wnt and Shh. Front Cell Dev Biol 9:758203.
70. Seppala, M., G.J. Fraser, A.A. Birjandi, et al. (2017) Sonic Hedgehog Signaling and Development of the Dentition. J Dev Biol 5.
71. Mogollón, I., J.E. Moustakas-Verho, M. Niittykoski, et al. (2021) The initiation knot is a signaling center required for molar tooth development. Development 148.
72. Bello, I.O., M.A. Alrabeeah, N.F. AlFouzan, et al. (2021) FAK, paxillin, and PI3K in ameloblastoma and adenomatoid odontogenic tumor. Clin Oral Investig 25:1559-1567.

73. Hanks, S.K., L. Ryzhova, N.Y. Shin, et al. (2003) Focal adhesion kinase signaling activities and their implications in the control of cell survival and motility. Front Biosci 8:d982-96. 10.2741/1114
74. Eapen, A., A. Ramachandran, and A. George (2012) Dentin phosphoprotein (DPP) activates integrin-mediated anchorage-dependent signals in undifferentiated mesenchymal cells. J Biol Chem 287:5211-24.
75. Park, J.H., J.H. Yoon, Y.S. Lim, et al. (2010) TAT-Hsp27 promotes adhesion and migration of murine dental papilla-derived MDPC-23 cells through beta1 integrin-mediated signaling. Int J Mol Med 26:373-8.
76. Premaraj, S., I. Souza, and T. Premaraj (2013) Focal adhesion kinase mediates β-catenin signaling in periodontal ligament cells. Biochem Biophys Res Commun 439:487-92.
77. Lee, J.M., C. Qin, O.H. Chai, et al. (2022) MSX1 Drives Tooth Morphogenesis Through Controlling Wnt Signaling Activity. J Dent Res 101:832-839. 10.1177/00220345211070583
78. Yamada, S., R. Lav, J. Li, et al. (2019) Molar Bud-to-Cap Transition Is Proliferation Independent. J Dent Res 98:1253-1261. 10.1177/0022034519869307
79. Duan, X. (2014) Ion channels, channelopathies, and tooth formation. J Dent Res 93:117-25. 10.1177/0022034513507066
80. Kantaputra, P., A. Butali, S. Eliason, et al. (2023) CACNA1S mutation-associated dental anomalies: A calcium channelopathy. Oral Dis.
81. Laugel-Haushalter, V., S. Morkmued, C. Stoetzel, et al. (2018) Genetic Evidence Supporting the Role of the Calcium Channel, CACNA1S, in Tooth Cusp and Root Patterning. Front Physiol 9:1329.
82. Stoneking, M. and F. Delfin (2010) The human genetic history of East Asia: weaving a complex tapestry. Curr Biol 20:R188-93. 10.1016/j.cub.2009.11.052

ABSTRACT (KOREAN)

몽골인들에게서 비증후군 (non-syndromic) 치아결손와 관련된 유전자 변이 분석

연세대학교 대학원 치의학과

<지도교수 **이 재 훈**>

셈지드 데지드너러브

치아의 결손을 뜻하는 TA (tooth agenesis)는 흔하게 발견되는 발달이상이다. TA와 연관된 다수 유전자들의 발견은 치아 발달에 대한 이해도를 높이는 데에 기여하고 있다. 그러나 원인이 되는 유전자 및 관련된 신호 전달 경로는 여전히 명확히 밝혀지지 않았다. 본 연구는 WES (whole-exome sequencing)와 생물정보학 분석을 이용해 아홉 몽골 가족들에게서 비증후군 (non-syndromic) TA와 관련된 유전자 변이를 찾는 것을 목표로 하였다. TA가 유전성을 띈 세 가족과 유전성을 띄지 않은 여섯 가족을 포함한 총 41명의 지원자가 연구에

포함되었다. 그 중 비증후군 TA 환자 14명의 타액 샘플에 대해 WES 분석을 수행했다. 후보 유전자는 변이 필터링과 분리 분석을 통해 선별되었다. 필터링 된 변이들은 in silico 돌연변이 영향 분석을 진행했다.

우리는 WES 분석을 통해 TA와 연관된 21개의 변이를 발견하였고, 이중 5개가 모든 필터링 기준을 충족하였다. 해당 변이는 MAST4, ITGA6, PITX2, CACNA1S, CDON 유전자의 엑손 (exon) 영역에 위치했다. PITX2의 변이는 가족의 TA 유전성에 상관없이 8명의 환자에게서 발견된 반면, MAST4의 변이는 유전성을 띈 가족의 환자 6명에게서 발견되었다. 본 연구는 여러 가족 그룹에서 TA와 관련된 다양한 유전자 변이 후보를 확인하였고, 그 중 PITX2의 변이가 가장 일반적으로 발견되었다.

우리 연구 결과는 Wnt 신호 전달 경로와의 연관되어 있는 MAST4를 TA 발병에 대한 새로운 후보 유전자로 제시한다. 또한, 우리는 5가지 후보 유전자가 치아 발달에 필수적인 기능인 초점 부착 (focal adhesion), 칼슘 채널 복합체 (calcium channel complex)와 두드러지게 관련되어 있음을 발견하였다.

TA와 관련된 새로운 병원성 유전자의 식별은 질병을 일으키는 분자 메커니즘에 대한 이해를 향상시켜 더 나은 진단, 예방 및 치료를 가능케한다. 바이오마커를 기반으로 한 TA의 조기 발견은 치아 관리를 개선하고 교정 및 보철 치료를 용이하게 할 수 있다.

핵심 되는 말: 치아 결손, 유전자 변이, 전장엑솜시퀀싱, 생물정보학 분석, 인실리코 돌연변이 분석, 몽골 집단