

DATA REPORT

Whole-exome sequencing analysis of supernumerary teeth occurrence in Japanese individuals

Masahiro Takahashi^{1,5}, Kazuyoshi Hosomichi^{2,5}, Tetsutaro Yamaguchi¹, Keisuke Yano³, Takahiro Funatsu⁴, Mohamed Adel¹, Shugo Haga¹, Koutaro Maki¹ and Atsushi Tajima²

A common disorder of human dentition is the existence of supernumerary teeth. Impacted supernumerary teeth occur most frequently in the maxillary incisor area and are termed mesiodens. We conducted whole-exome sequencing of non-syndromic Japanese individuals possessing supernumerary teeth to identify genes and/or loci involved in the pathogenesis of the condition.

Human Genome Variation (2017) 4, 16046; doi:10.1038/hgv.2016.46; published online 26 January 2017

Teeth are an important vertebrate organ, especially in humans because of their use in feeding, mastication, and speech. Tooth development and formation involve various complex processes, including epithelial-mesenchymal transitions, that are precisely controlled by molecular signaling pathways at each developmental stage.¹ A common disorder found in human dentition is the existence of supernumerary teeth. There are 20 teeth in the primary dentition and 28 teeth, excluding third molars, in the permanent dentition. Supernumerary teeth are defined as extra teeth—more than the normal human complement.² The frequency of this condition in permanent dentition ranges from approximately 0.1 to 3.6%, with varied ranges among populations,³ and is more prevalent among males than females, with a ratio of approximately 2:1.⁴ Supernumerary teeth are classified by their various morphologies or location. The most common supernumerary teeth are located in the maxillary incisor area and are called mesiodens.⁵ Mesiodens are typically conical supernumerary teeth that are small and peg-shaped. The autosomal-dominant transmission pattern of non-syndromic multiple supernumerary teeth has been reported.⁶

Cleidocranial dysplasia (CCD) is a human syndrome that presents with supernumerary teeth. It is an autosomal-dominant disorder characterized by short stature, hypoplasia or aplasia of clavicles, delayed closure of the cranial fontanelles and sutures, maxillary hypoplasia, Wormian bones, frontal bossing, wide pubic symphysis, and dental abnormalities including supernumerary teeth.⁷ Mutations in runt related transcription factor 2 (*RUNX2*), an essential transcription factor for osteoblast differentiation and skeletal development⁸, have been identified as a main cause for CCD.⁹ In addition to CCD, other human syndromes are also associated with supernumerary teeth.¹⁰

Various mouse models have been used to explore the etiology of supernumerary teeth and to identify candidate genes involved in the pathogenesis of this condition. Supernumerary teeth in the diastema region in front of the first molars are exhibited by *Sprouty2* and *Sprouty4* null mice (fibroblast growth factor (*Fgf*)

antagonists), Tabby mice, which have a mutation in ectodysplasin (*Eda*), transgenic mice overexpressing *Eda* or ectodysplasin receptor (*Edar*), paired box 6 (*Pax6*) mutant mice, and growth arrest specific 1 (*Gas1*) null mutant mice. Furthermore, supernumerary teeth in both incisor and molar regions are exhibited by specificity protein 6 (*Sp6*) deficient mice and sclerostin domain containing 1 (*Sostdc1*) (also known as *ectodin*, *Wise* and *USAG-1*) null mutant mice.¹⁰ Mice possessing conditional knockout of the adenomatous polyposis coli (*Apc*) gene under the control of the human keratin 14 (*K14*) promoter also exhibited supernumerary teeth. This study indicated that the occurrence of supernumerary teeth by loss of *Apc* was through activation of Wnt/ β -catenin signaling.¹¹ The normal number of teeth in mice is, however, less than that of humans, and their teeth are not replaced because they have only primary dentition. Therefore, mouse models may not be the most appropriate ones for determining the etiology of supernumerary teeth in humans.

In the present study, we selected four Japanese families, Families A, B, C, and D. In three families (Families A, B and D), both first and second generations have supernumerary teeth. We do not know whether the first generation in Family C is affected. However, the second generation in Family C has impacted supernumerary teeth in the maxillary incisor area. Most affected individuals have one supernumerary tooth; however, II-2 in Family A has two (Figure 1a, b). The occurrence of supernumerary teeth in most of these families indicates an autosomal-dominant pattern of inheritance. All subjects were diagnosed by examining a panoramic radiograph or upper anterior occlusal radiographs and interviewing the patient about their medical and dental history. All subjects were free of any syndrome or congenital anomaly such as cleft lip or palate. All subjects provided written informed consent and the study was approved by the Ethics Committee of Showa and Kanazawa University.

Saliva samples were collected from 16 individuals (10 affected and 6 unaffected). DNA was extracted from the samples using Oragene DNA (DNA Genotek, Ottawa, Canada) and subjected to

¹Department of Orthodontics, School of Dentistry, Showa University, Tokyo, Japan; ²Department of Bioinformatics and Genomics, Graduate School of Advanced Preventive Medical Sciences, Kanazawa University, Ishikawa, Japan; ³Verde Orthodontic Dental Clinic, Tokyo, Japan and ⁴Division of Dentistry for Persons with Disabilities, Department of Special Needs Dentistry, School of Dentistry, Showa University, Tokyo, Japan.

Correspondence: T Yamaguchi (tyamaguchi@dent.showa-u.ac.jp)

⁵These authors contributed equally to this work.

Received 7 October 2016; revised 18 November 2016; accepted 20 November 2016

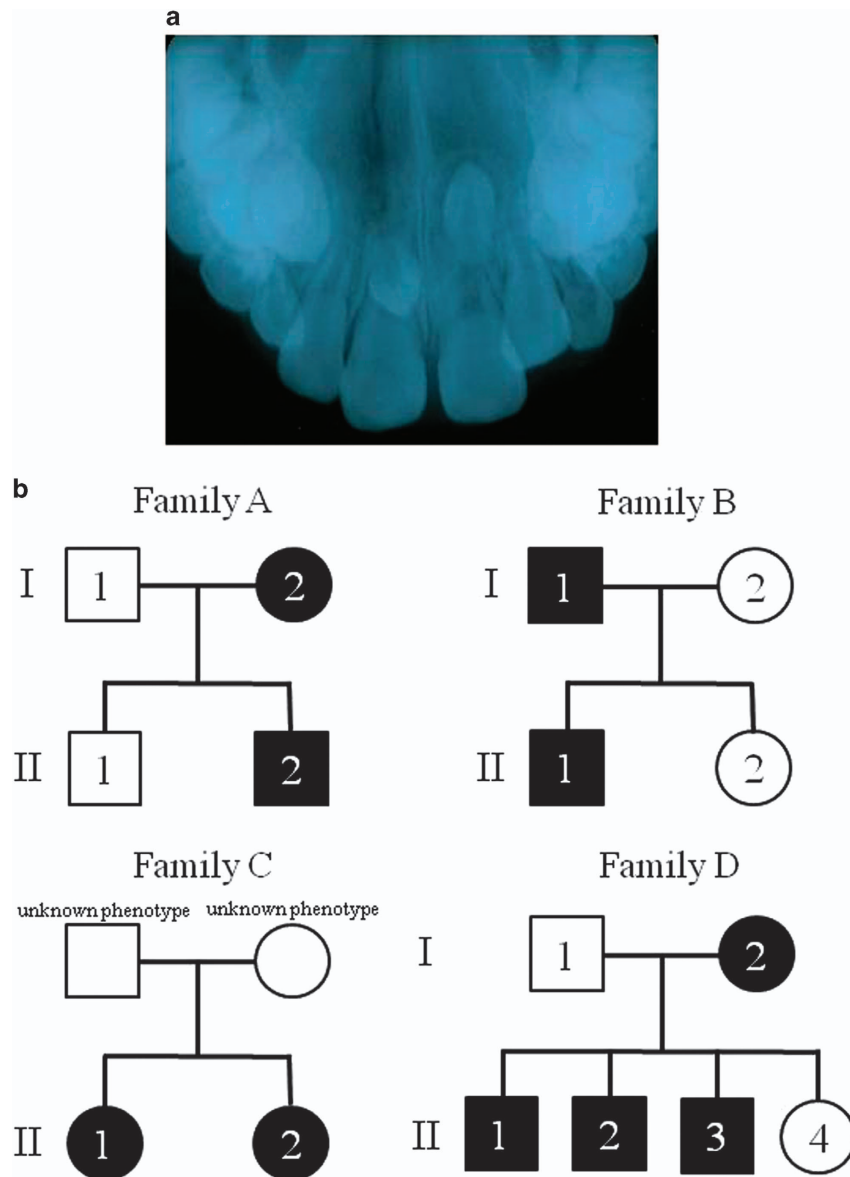


Figure 1. (a) Upper anterior occlusal radiograph of II-2 in Family A with two impacted supernumerary teeth in the maxillary incisor area. (b) Family pedigrees of Japanese individuals with supernumerary teeth. Squares and circles denote males and females, respectively. Filled symbols indicate affected individuals. The phenotype of the first generation in Family C is unknown.

whole-exome sequencing. DNA samples (3 µg) were subjected to exome capture using the Sure-Select Human All Exon Kit (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer's instructions. The captured DNA underwent high-throughput sequencing using the HiSeq2500 system (Illumina, San Diego, CA, USA). All prepared flow cells were run on the Illumina HiSeq2500 using paired-end 100-bp reads. Reads were mapped to the reference genome (UCSC hg19) using the Burrows-Wheeler Aligner (BWA) v.0.7.9.¹² BWA-generated SAM files were sorted and indexed using SAMtools v.0.1.18.¹³ The files obtained in BAM format were analyzed using GATK v.2.7 following their best practice guidelines.¹⁴ To perform gene annotation, we used the RefSeq gene database (build hg19),¹⁵ while variant annotation was based on the dbSNP (dbSNP 135), the 1000 Genomes Project database,¹⁶ and data of 1,208 Japanese individuals from the Human Genetic Variation database (HGVD) (<http://www.genome.med.kyoto-u.ac.jp/SnpDB/index.html>). In all family samples, variants detected in exome sequencing data were further analyzed by performing three filtering steps based on

different criteria. Rare variants with frequencies < 3% in the Japanese population were selected from HGVD to obtain rare variants as potential candidates.

Whole-exome sequencing identified 149 single-nucleotide variants (SNVs) and seven small insertions and deletions (indels) that co-segregated in Family A in the fully penetrant autosomal-dominant mode of inheritance. Moreover, 157 SNVs and seven indels co-segregated in Family B, 407 SNVs and 28 indels co-segregated in Family C, and 16 SNVs and four indels co-segregated in Family D. However, we did not identify any common gene variants in all four families. Three families (Families A, C, and D) all had variants in SCO-spondin (*SSPO*) (Table 1). It is known that *SSPO* plays a role in central nervous system development.^{17,18} We detected 23 common gene variants in two families. Families A and B both had variants in EF-hand calcium binding domain 5 (*EFCAB5*), testis expressed 15 (*TEX15*), and tumor suppressor candidate 1 (*TUSC1*). Families A and C both had variants in absent in melanoma 1-like (*AIM1L*), cadherin 26 (*CDH26*), exocyst complex component 3-like 4 (*EXOC3L4*), family

Table 1. Summary of candidate mutations in SSPO identified by whole-exome sequencing

	Chr	Position	Ref	Obs	Function	Gene	Exonic function	AACchange	HGVD	1,000 g	dbSNP135	
Family A	7	149493577	A	T	Exonic	SSPO	Nonsynonymous SNV	NM_198455 c.A6653T	p.Q2218L	0.02193	0.04	rs73727613
Family C	7	149484985	C	A	Exonic	SSPO	Stopgain SNV	NM_198455 c.C3740A	p.S1247X	0.022665	0.0041	rs146309075
Family C	7	149504070	G	A	Exonic	SSPO	Nonsynonymous SNV	NM_198455 c.G8894A	p.C2965Y	0.021668	0.01	rs140118386
Family D	7	149500107	G	A	Exonic	SSPO	Nonsynonymous SNV	NM_198455 c.G7733A	p.R2578H	0.004177	0	

Abbreviations: AACchange, amino acid change; dbSNP135, database single-nucleotide polymorphisms 135; HGVD, Human Genetic Variation Database; Obs, observed base; Ref, reference base; SNV, single-nucleotide variant; SSPO, SCO-spondin; 1000 g, 1000 Genomes Project database.

Table 2. Genes with variants detected in two families

Gene	Presence of gene variants in families			
	Family A	Family B	Family C	Family D
1 Absent in melanoma 1-like (<i>AIM1L</i>)	○		○	
2 Agrin (<i>AGRN</i>)		○	○	
3 Ataxin 1 (<i>ATXN1</i>)		○	○	
4 Cadherin 26 (<i>CDH26</i>)	○		○	
5 Chromosome 21 open reading frame 58 (<i>C21orf58</i>)		○	○	
6 Complement factor B (<i>CFB</i>)		○	○	
7 EF-hand calcium binding domain 5 (<i>EFCAB5</i>)	○	○		
8 Exocyst complex component 3 like 4 (<i>EXOC3L4</i>)	○		○	
9 Family with sequence similarity 186 member A (<i>FAM186A</i>)	○		○	
10 Fanconi anemia complementation group E (<i>FANCE</i>)		○	○	
11 Formin like 1 (<i>FMNL1</i>)		○	○	
12 FXYD domain containing ion transport regulator 4 (<i>FXYD4</i>)	○		○	
13 Hemicentin 1 (<i>HMCN1</i>)		○	○	
14 Immunoglobulin superfamily member 9B (<i>IGSF9B</i>)		○	○	
15 <i>KIAA1614</i>		○	○	
16 <i>LOC100652824</i>	○		○	
17 MGA, MAX dimerization protein (<i>MGA</i>)		○	○	
18 Phospholipase C eta 2 (<i>PLCH2</i>)		○	○	
19 Polycystin 1 like 2 (<i>PKD1L2</i>)	○		○	
20 Ring finger protein 207 (<i>RNF207</i>)		○	○	
21 Testis expressed 15 (<i>TEX15</i>)	○	○		
22 Transketolase-like 1 (<i>TKTL1</i>)	○		○	
23 Tumor suppressor candidate 1 (<i>TUSC1</i>)	○	○		

with sequence similarity 186 member A (*FAM186A*), FXYD domain-containing ion transport regulator 4 (*FXYD4*), polycystin 1 like 2 (*PKD1L2*), *LOC100652824*, and transketolase-like 1 (*TKTL1*). Families B and C both had variants in agrin (*AGRN*), ataxin 1 (*ATXN1*), chromosome 21 open reading frame 58 (*C21orf58*), complement factor B (*CFB*), Fanconi anemia complementation group E (*FANCE*), formin-like 1 (*FMNL1*), hemicentin 1 (*HMCN1*), immunoglobulin superfamily member 9B (*IGSF9B*), *KIAA1614*, MGA, MAX dimerization protein (*MGA*), phospholipase C eta 2 (*PLCH2*), and ring finger protein 207 (*RNF207*) (Table 2). Up to date, none of these genes are known to have a function in tooth development.

This is the first study to demonstrate the genetic landscape in non-syndromic Japanese patients with impacted supernumerary teeth in the maxillary incisor area. Whole-exome sequencing is an effective strategy to elucidate the etiology of dental diseases, and our findings provide a basis for further exploration of the pathological mechanisms of supernumerary teeth.

HGV DATABASE

The relevant data from this Data Report are hosted at the Human Genome Variation Database at <http://dx.doi.org/10.6084/m9.figshare.hgv.938>; <http://dx.doi.org/10.6084/m9.figshare.hgv.941>; <http://dx.doi.org/10.6084/m9.figshare.hgv.944>; <http://dx.doi.org/10.6084/m9.figshare.hgv.947>

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI (16K15843; Grant-in-Aid for challenging Exploratory Research to TY) and by the Ministry of Education, Culture, Sports, Science and Technology (MEXT)—Supported Program for the Strategic Research Foundation at Private Universities, 2012–2017.

COMPETING INTERESTS

The authors declare no conflict of interest.

REFERENCES

- Marianna B. Molecular genetics of tooth development. *Curr Opin Genet Dev* 2009; **19**: 504–510.
- Schulze C. Developmental abnormalities of the teeth and jaws. In: Gorlin RJ, Goldman HM (eds) *Thomas' Oral Pathology* 6th edn, Vol. 1. Mosby: St Louis, USA, 1970, pp 96–183.
- Yusof WZ. Non-syndromic multiple supernumerary teeth: literature review. *J Can Dent Assoc* 1990; **56**: 147–149.
- Rajab LD, Hamdan MA. Supernumerary teeth: review of the literature and a survey of 152 cases. *Int J Paediatr Dent* 2002; **12**: 244–254.
- Hyun HK, Lee SJ, Lee SH, Hahn SH, Kim JW. Clinical characteristics and complications associated with mesiodentia. *J Oral Maxillofac Surg* 2009; **67**: 2639–2643.
- Batra P, Duggal R, Parkash H. Non-syndromic multiple supernumerary teeth transmitted as an autosomal dominant trait. *J Oral Pathol Med* 2005; **34**: 621–625.
- Mundlos S. Cleidocranial dysplasia: clinical and molecular genetics. *J Med Genet* 1999; **36**: 177–182.

- 8 Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki K, Deguchi K *et al*. Targeted disruption of *Cbfa1* results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell* 1997; **89**: 755–764.
- 9 Mundlos S, Otto F, Mundlos C, Mulliken JB, Aylsworth AS, Albright S *et al*. Mutations involving the transcription factor *CBFA1* cause cleidocranial dysplasia. *Cell* 1997; **89**: 773–779.
- 10 Wang XP, Fan J. Molecular genetics of supernumerary tooth formation. *Genesis* 2011; **49**: 261–277.
- 11 Wang XP, O'Connell DJ, Lund JJ, Saadi I, Kuraguchi M, Turbe-Doan A *et al*. *Apc* inhibition of *Wnt* signaling regulates supernumerary tooth formation during embryogenesis and throughout adulthood. *Development* 2009; **136**: 1939–1949.
- 12 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**: 1754–1760.
- 13 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
- 14 DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C *et al*. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011; **43**: 491–498.
- 15 Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O *et al*. RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res* 2014; **42**: D756–D763.
- 16 Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE *et al*. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**: 56–65.
- 17 Gobron S, Creveaux I, Meiniel R, Didier R, Herbet A, Bamdad M *et al*. Subcommissural organ/Reissner's fiber complex: characterization of SCO-spondin, a glycoprotein with potent activity on neurite outgrowth. *Glia* 2000; **32**: 177–191.
- 18 Meiniel A. SCO-spondin, a glycoprotein of the subcommissural organ/Reissner's fiber complex: evidence of a potent activity on neuronal development in primary cell cultures. *Microsc Res Tech* 2001; **52**: 484–495.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

© The Author(s) 2017