Data in Brief 32 (2020) 106224

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Dataset on amelogenesis-related genes variants (*ENAM* and *ENAM* interacting genes) and on human leukocyte antigen alleles (DQ2 and DQ8) distribution in children with and without molar-incisor hypomineralisation (MIH)

Luka Hočevar[a], Jernej Kovač[b], Katarina Trebušak Podkrajšek[b,c], Saba Battelino[d,e], Alenka Pavlič[a,f,*]

[a] *Department of Paediatric and Preventive Dentistry, Faculty of Medicine, University of Ljubljana, Hrvatski trg 6, Ljubljana 1000, Slovenia*
[b] *University Children's Hospital, University Medical Centre Ljubljana, Bohoričeva 20, Ljubljana, Slovenia*
[c] *Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, Ljubljana, Slovenia*
[d] *Department of Otorhinolaryngology and Cervicofacial Surgery, University Medical Centre Ljubljana, Zaloška 2, Ljubljana, Slovenia*
[e] *Department of Otorhinolaryngology, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, Ljubljana, Slovenia*
[f] *Department of Paediatric and Preventive Dentistry, University Medical Centre Ljubljana, Hrvatski trg 6, Ljubljana, Slovenia*

ARTICLE INFO

Article history:
Received 30 July 2020
Revised 20 August 2020
Accepted 21 August 2020
Available online 25 August 2020

ABSTRACT

All children, who were born in 2004 and had undergone surgical treatment for recurrent acute tonsillitis and/or acute otitis media at the ear, nose and throat clinic (ENT) between 2004 and 2010, were called on dental examination and blood sampling. Out of 441 invitees, 113 children and their parents/legal guardians agreed to participate. The following data

Abbreviations: MIH, molar-incisor hypomineralisation; ENT, ear, nose and throat clinic; HLA, human leukocyte antigen; SNP, single nucleotide polymorphism; HRM, high resolution melting; FPM, First permanent molar; PEB, Post-eruptive enamel breakdown; CADD, Combined annotation dependent depletion.

DOI of original article: 10.1016/j.archoralbio.2020.104848

* Corresponding author at: Department of Paediatric and Preventive Dentistry, Faculty of Medicine, University of Ljubljana, Hrvatski trg 6, Ljubljana 1000, Slovenia.
E-mail address: alenka.pavlic@mf.uni-lj.si (A. Pavlič).

https://doi.org/10.1016/j.dib.2020.106224

Keywords:
Molar-incisor hypomineralisation
Aetiology
Human leukocyte antigen
The amelogenesis-related genes
Single nucleotide polymorphism

from this group of subjects are presented: the presence of clinical signs of molar–incisor hypomineralisation (MIH), the distribution of human leukocyte antigen (HLA) alleles DQ2 and DQ8 and eight single nucleotide polymorphisms (SNPs) located in amelogenesis-related genes (rs3796704 in the *ENAM* gene, rs546778141 in the *AMBN* gene, rs2106416 in the *AMELX* gene, rs7660807 and rs35286445 in the *AMTN* gene, rs4870723 in the *COL14A1* gene, rs2245803 in the *MMP20* gene, and rs3828054 in the *TUFT1* gene). Data on clinical signs of MIH were collected in accordance with the recommendation and on the proposed MIH clinical data recording sheet [1], and with appropriate preliminary training and calibration. Data on HLA DQ2 and DQ8 haplotypes and on SNPs of amelogenesis-related genes were obtained using DNA isolated from blood samples taken from subjects. The HLA DQ2 and DQ8 alleles were determined using the EliGene® Coeliac RT Kits (90,048-RT; Elisabeth Pharmacon spol. s.r.o., Brno-Židenice, Czech Republic) on a 7500 Fast RT-PCR System (Applied Biosystems, Waltham, MA, USA). The distributions of SNPs in the amelogenesis-related genes were determined using high resolution melting (HRM) using the Type-IT HRM Master Mix (Qiagen), TaqMan genotyping assays (ID: C__25766207_10; Thermo Fisher Scientific, Waltham, MA, USA) with the TaqMan Universal Master Mix II, or Sanger sequencing using sequencing master mix BigDye® Terminator v3.1 (Applied Biosystems) and ABI 3500 Genetic Analyser (Applied Biosystems).

L. Hočevar, J. Kovač, K. Trebušak Podkrajšek, S. Battelino, A. Pavlič, 2020. The possible influence of genetic aetiological factors on molar–incisor hypomineralisation, Arch. Oral. Biol. 118, 104848. https://doi.org/10.1016/j.archoralbio.2020.104848.

Specifications Table

Subject	Medicine and Dentistry (General)
Specific subject area	Molar-incisor hypomineralisation (MIH) aetiology remains unclear, but might develop due to a simultaneous or additive effect of environmental and genetic factors.
Type of data	Tables, figure and raw DNA data.
How data were acquired	Clinical signs of MIH were diagnosed by dental examination performed in accordance with recommendation [1]. From the isolated DNA samples, presence of the alleles HLA DQ2 and HLA DQ8 was determined using 7500 Fast RT-PCR System (Applied Biosystems, Waltham, MA, USA). SNPs in *ENAM* were detected using 7500 Fast RT-PCR System (Applied Biosystems) and analysed with high resolution melting (HRM) version 2.0.1 software (Applied Biosystems) or 7500 software version 2.0.6 (Applied Biosystems). SNPs in the *ENAM* interacting genes were identified using online String database [2,3,4] software and the combined annotation dependent depletion [5] algorithm and later on determined with Sanger sequencing using ABI Genetic Analyser 3500 (Applied Biosystems) and analysed with the Sequencing Analysis Software v5.4. (Applied Biosystems).

(*continued on next page*)

Data format	Raw and analysed data.
Parameters for data collection	Dental examinations were conducted by one trained and calibrated examiner (L.H.) under standardised conditions (on the dental chair and under artificial light using a dental mirror and a probe). DNA was isolated from a blood sample of each subject using a FlexiGene DNA isolation kit (Qiagen, Hilden, Germany) according to manufacturer's instructions.
Description of data collection	MIH was recorded using the proposed MIH clinical data recording sheet (long form) [1]. HLA DQ2 and HLA DQ8 alleles were determined using the EliGene® Coeliac RT Kits (Elisabeth Pharmacon spol. s.r.o., Brno-Židenice, Czech Republic) according to the producer's recommendations. SNPs in *ENAM* were analysed using HRM analysis or predesigned TaqMan genotyping assays (Applied Biosystems). SNPs in the *ENAM* interacting genes were determined by Sanger sequencing.
Data source location	Institution: University Medical Centre Ljubljana City: Ljubljana Country: Slovenia Latitude and longitude (GPS): 46.052339, 14.520787 (46°03′08.4″N, 14°31′14.8″E)
Data accessibility	The data is available with this article.
Related research article	L. Hočevar, J. Kovač, K. Trebušak Podkrajšek, S. Battelino, A. Pavlič, 2020. The possible influence of genetic aetiological factors on molar–incisor hypomineralisation, Arch. Oral. Biol. 118, 104848. https://doi.org/10.1016/j.archoralbio.2020.104848.

Value of the Data

- The dataset can be used for further analysis of genetic polymorphisms in different *ENAM* interacting genes and their possible influence on development of molar-incisor hypomineralisation (MIH).
- The dataset provides possibilities for further analysis that could help clarify the possible genetic involvement in the development of MIH.
- The present data provide further insight and possibilities for analyses of diagnosed MIH phenotypes and identified genotypes.

1. Data Description

Fig. 1 represents the distribution of MIH lesions in individual groups of first permanent molars (FPM) and incisor teeth in the studied cohort. Lesions detected were demarcated opacities, post-eruptive enamel breakdown (PEB) and atypical restorations. Fig. 1 represents analysed data from the Supplementary Table 1.

Fig. 2 shows the distribution of the clinical status of MIH signs of affected FPMs and incisors in the studied cohort. Individual groups of teeth were noted using FDI two-digit notation. Clinical status included white or creamy demarcated opacities, yellow or brown demarcated opacities, PEB and atypical restorations. Fig. 2 represents analysed data from the Supplementary Table 1.

Fig. 3 represents interacting map from online String database [2] software accessed on September 19th 2018. It was used for identification of genes interacting with *ENAM* at the time of the study. The genes that interacted with *ENAM* were *PARL, MTA2, COL14A1, BMP3, MMP20, AMELX, AMBN, TUFT1, AMTN*, and *AMELY*. Only SNPs in *ENAM* interacting genes that met the inclusion criteria described in 2.4. (SNP selection) were analysed.

Table 1 shows raw data on HLA DQ2 and DQ8 alleles distribution, genotype and distribution of eight SNPs in *ENAM* (rs3796704) and *ENAM* interacting genes (rs546778141 in the *AMBN* gene, rs2106416 in the *AMELX* gene, rs7660807 and rs35286445 in the *AMTN* gene, rs4870723 in the *COL14A1* gene, rs2245803 in the *MMP20* gene, and rs3828054 in the *TUFT1* gene), with reference and alternative allele.

Table 1
The Human leukocyte antigen (HLA) DQ2 and DQ8 alleles distribution, genotype and distribution of eight SNPs in *ENAM* and *ENAM* interacting genes in the studied cohort.

			Genotypes (reference allele/alternative allele)							
	HLA		***AMBN***	***AMELX***	***AMTN***	***AMTN***	***COL14A1***	***ENAM***	***MMP20***	***TUFT1***
ID	**DQ2**	**DQ8**	**rs546778141**	**rs2106416**	**rs7660807**	**rs35286445**	**rs4870723**	**rs3796704**	**rs2245803**	**rs3828054**
1	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AG (A/G)
2	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
3	–	+	GG (G/A)	TT (C/T)	AT (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
4	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AG (A/G)	TG (T/G)	AA (A/G)
5	–	–	GG (G/A)	TT (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
6	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	CC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
7	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	CC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
8	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
9	+	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AG (A/G)	GG (T/G)	AA (A/G)
10	+	–	GG (G/A)	TT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
11	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
12	+	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
13	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
14	–	–	GG (G/A)	TT (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
15	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	CC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
16	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	CC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
17	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
18	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	CC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
19	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AG (A/G)
20	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	CC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
21	–	–	GG (G/A)	TT (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AG (A/G)
22	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
23	+	–	GG (G/A)	TT (C/T)	AT (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
24	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
25	+	+	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	CC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
26	+	–	GG (G/A)	TT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
27	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
28	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	CC (A/C)	AA (A/G)	GG (T/G)	AG (A/G)
29	–	+	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
30	+	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	GG (A/G)
31	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AG (A/G)
32	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AG (A/G)
33	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
34	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
35	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
36	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AG (A/G)
37	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
38	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
39	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
40	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
41	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
42	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AG (A/G)
43	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AG (A/G)
44	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
45	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
46	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
47	+	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
48	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
49	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AG (A/G)
50	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
51	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AG (A/G)
52	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
53	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
54	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
55	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
56	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AG (A/G)	TG (T/G)	AA (A/G)
57	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)

(*continued on next page*)

Table 1 (*continued*)

			Genotypes (reference allele/alternative allele)							
	HLA		***AMBN***	***AMELX***	***AMTN***	***AMTN***	***COL14A1***	***ENAM***	***MMP20***	***TUFT1***
ID	**DQ2**	**DQ8**	**rs546778141**	**rs2106416**	**rs7660807**	**rs35286445**	**rs4870723**	**rs3796704**	**rs2245803**	**rs3828054**
58	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AG (A/G)	TG (T/G)	GG (A/G)
59	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
60	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
61	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
62	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AG (A/G)
63	–	–	GG (G/A)	TT (C/T)	AT (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
64	–	+	GG (G/A)	TT (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
65	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
66	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
67	–	–	GG (G/A)	CC (C/T)	AT (A/T)	CT (C/T)	CC (A/C)	AG (A/G)	TT (T/G)	AA (A/G)
68	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	CC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
69	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AG (A/G)	GG (T/G)	AA (A/G)
70	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
71	+	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
72	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AG (A/G)
73	+	+	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AG (A/G)
74	+	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
75	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
76	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
77	–	–	GG (G/A)	TT (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
78	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
79	–	+	GG (G/A)	TT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
80	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AG (A/G)
81	–	–	GG (G/A)	TT (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
82	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
83	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
84	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
85	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
86	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
87	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	TT (T/G)	AG (A/G)
88	–	+	GG (G/A)	CC (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
89	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
90	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
91	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
92	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AG (A/G)
93	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
94	–	–	GG (G/A)	TT (C/T)	TT (A/T)	CC (C/T)	AA (A/C)	AG (A/G)	TG (T/G)	AA (A/G)
95	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
96	–	–	GG (G/A)	CC (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
97	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
98	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
99	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
100	–	–	GG (G/A)	TT (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
101	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
102	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
103	–	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
104	+	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AG (A/G)
105	–	–	GG (G/A)	CC (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
106	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
107	–	+	GG (G/A)	CC (C/T)	AA (A/T)	CT (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)
108	–	+	GG (G/A)	CT (C/T)	AA (A/T)	CT (C/T)	AC (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
109	+	–	GG (G/A)	CC (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TG (T/G)	AG (A/G)
110	+	–	GG (G/A)	CT (C/T)	AT (A/T)	CC (C/T)	AC (A/C)	AA (A/G)	TT (T/G)	AA (A/G)
111	+	–	GG (G/A)	CC (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	GG (T/G)	AA (A/G)
112	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AC (A/C)	AG (A/G)	TG (T/G)	AA (A/G)
113	–	–	GG (G/A)	CT (C/T)	AA (A/T)	CC (C/T)	AA (A/C)	AA (A/G)	TG (T/G)	AA (A/G)

HLA, human leukocyte antigen; A, Adenine; C, Cytosine; G, Guanine; T, Thymine.

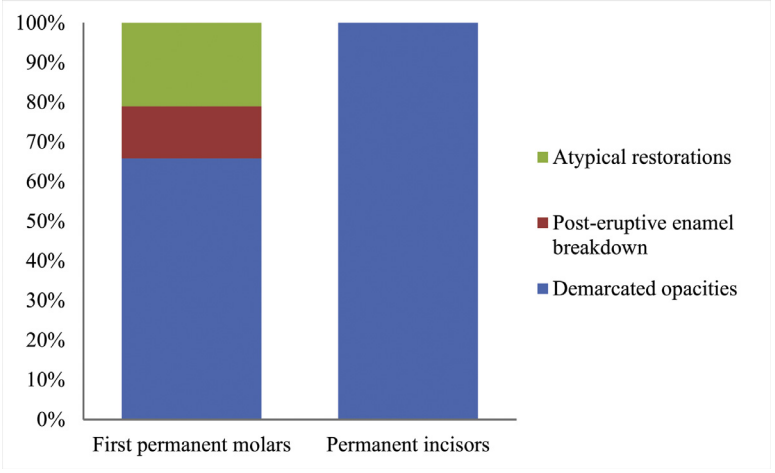

Fig. 1. Distribution of molar–incisor hypomineralisation lesions.

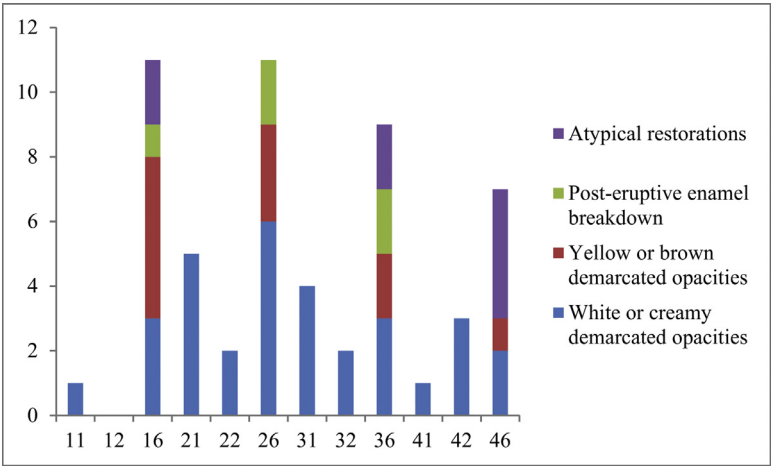

Fig. 2. Distribution of the clinical status of molar–incisor hypomineralisation (MIH) lesions among the individual groups of teeth using FDI two-digit notation in a group of MIH-affected subjects ($n = 22$). (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

Supplementary Table 1 shows raw data on the distribution of enamel developmental lesions recorded on the labial/buccal, incisal/occlusal and lingual/palatal surfaces of each tooth in individuals diagnosed with MIH ($n = 22$), as recommended by Ghanim et al. [1]. In accordance with this recommendation, the presence of lesion, its status and extension for each of these tooth surfaces were recorded on a MIH clinical data recording sheet (long form) [1]. Only modification we made to the MIH clinical data recording sheet was in omitting the description of eruption status of incisor and FPM teeth. Please note that the ID number of each individual is the same as in Table 1.

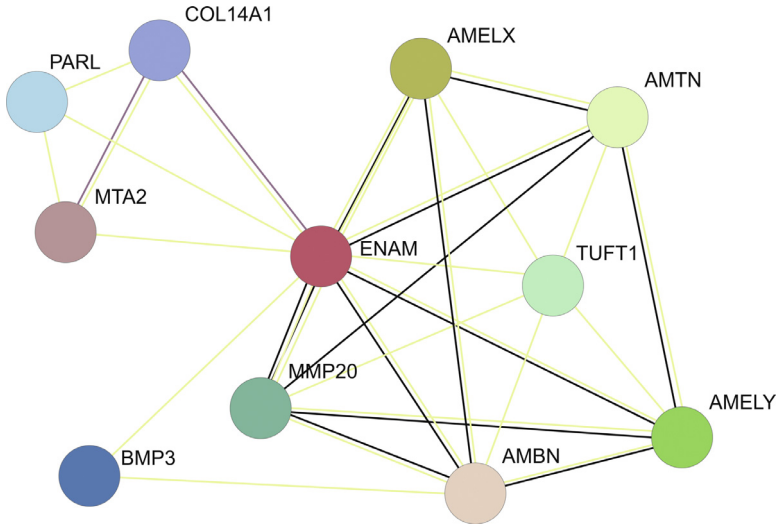

Fig. 3. *ENAM* gene interacting map from online String database [2] software (accessed on September 19th 2018).

Supplementary Table 2 represents information on custom-designed primers for *ENAM* analysis and analysis of SNPs in *ENAM* interacting genes. For each custom-designed primer there is information about gene/SNP, position (exon), forward and reverse primer sequence with corresponding melting temperatures and amplicon length.

Supplementary Table 3 represents raw data of the HRM analysis of SNPs in *ENAM* interacting genes with assigned representative sequences of HRM variants.

Supplementary material contains the raw representative DNA sequences of variants detected using HRM analysis and determined using Sanger sequencing.

2. Experimental design, materials and methods

2.1. Study subjects

441 children, born in 2004 and referred to the ENT for surgery in general anaesthesia between 2004 and 2010, were invited to participate in the study. Among those who responded to the invitation, individuals with the following conditions were excluded: cleft lip and/or palate, bronchopulmonary dysplasia, Wolff-Parkinson-White syndrome, neurofibromatosis, nephrotic syndrome, and those whose ENT medical record was either absent or unavailable, or whose quantity of blood sample was insufficient for further analysis. Finally, a total of 113 children met the criteria. Written informed consent was obtained from parents or their legal guardians to allow their children to participate in all procedures associated with the study. The whole protocol was approved by the Slovenian National Medical Ethics Committee (65/05/14).

Dental examination was performed under standardised conditions (on the dental chair and under artificial light using dental mirror and a probe) by one trained and calibrated investigator (L.H.). Wet and previously cleaned teeth were inspected for MIH following the criteria recommended by the European Academy of Paediatric Dentistry [6]. Findings were recorded using the long form for MIH clinical data recording sheet (Supplementary Table 1) as recommended by Ghanim et al. [1]. After completing dental examinations, the children were assigned into two groups: in the first, children with at least one FPM affected with MIH, and in the second those with no MIH affected FPM. Distributions of enamel lesions are presented in Figs. 1 and 2.

2.2. Blood sample collection and DNA handling

Whole blood samples were collected from each participant using a plastic tube (Vacutainer®) containing EDTA solution and genomic DNA was isolated using the FlexiGene DNA isolation kits (Qiagen, Hilden, Germany) according to manufacturer's instructions.

2.3. Analysis of HLA DQ2 and HLA DQ8 haplotypes

The presence of HLA DQ2 and HLA DQ8 alleles in each DNA sample was determined using the EliGene® Coeliac RT Kit (90,048-RT; Elisabeth Pharmacon, Brno, Czech Republic) on the 7500 Fast RT-PCR System (Applied Biosystems, Waltham, MA, USA), and was performed with two replicates. When the replicated results differed, another dilution of the DNA sample was prepared and the analysis repeated. Each reaction consisted of 8 ng of DNA and 8 µL of Master Mix (CELI-DQ2 Mix, CELI-DQ8 Mix or CELI-DR4 Mix) (Elisabeth Pharmacon). The PCR protocol comprised polymerase activation at 95 °C for 3 min, followed by 40 cycles of denaturation at 95 °C for 15 s and annealing at 58 °C for 40 s when the emission signal is collected.

2.4. SNP selection

Eight SNPs in seven genes (Table 1) were selected for analysis according to previously reported SNP (rs3796704 *ENAM*) association with MIH [7] or according to the following criteria: 1) gene interacting with *ENAM* gene (Fig. 3) [2], 2) minor allele frequency (MAF) $> 10\%$, 3) substitution that results in a change of the amino acid sequence of the protein, and 4) combined annotation dependent depletion (CADD) values of ≥ 20 [8,9] (SNP rs2106416 in *AMELX* with CADD value of 19.97 was also included).

To identify the genes interacting with *ENAM* we created an interactive map (Fig. 3) using the online String database [2] software [3,4]. All reported SNPs in each interacting gene were identified using the ENSEMBL database [10] and later evaluated. SNPs with MAF $< 10\%$ and silent substitutions were excluded. Remaining SNPs were analysed by bioinformatics analysis using the CADD [5] algorithm to identify those that were likely to have the greatest effects. Finally, all remaining SNPs in the *ENAM* interacting genes with CADD values ≥ 19.97 were included in further analysis.

2.5. Custom-designed primer selection

The PCR primers for *ENAM* gene analysis and selected SNPs in *ENAM* interacting genes were custom-designed. Reference sequence for each amplicon was acquired from ENSEMBL database [10]. Optimal primers (Supplementary Table 2) were designed by the Primer3 [11] online tool and analysed for possible polymorphisms with the SNP check tool [12]. The specificity of the constructed primers was checked using primer-BLAST interface [13]. Each PCR reaction consisted of 8 µL of DNA, 10 µL GoTaq® G2 Green Master Mix (Promega, Madison, WI, USA), 1 µL forward primer and 1 µL reverse primer. The PCR protocol comprised polymerase activation at 95 °C for 2 min, followed by 30 cycles of denaturation at 95 °C for 30 s, annealing at 60 °C for 30 s, extension at 72 °C for 30 s and final extension at 72 °C for 5 min. All amplicons were evaluated using 2% agarose gel electrophoresis of the PCR amplicons.

2.6. ENAM *gene analysis*

HRM analysis was used to detect SNPs in *ENAM* gene. Using custom-designed primer pairs (Supplementary Table 2) all exons except the first and the last exon were amplified. HRM was

performed using the Type-IT HRM Master Mix (Qiagen) and the 7500 Fast RT-PCR System (Applied Biosystems). Each HRM reaction consisted of 10 ng of DNA, 5 µL of the Type-IT HRM Master Mix, 0.2 µL forward primer, 0.2 µL reverse primer and 2.6 µL of Nuclease-free water. The PCR protocol comprised polymerase activation at 95 °C for 5 min, followed by 50 cycles of denaturation at 95 °C for 10 s, annealing at 60 °C for 25 s, and extension at 72 °C for 15 s. Melting curves were recorded between 70 °C and 98 °C with a 1% temperature ramp, and were aligned and normalised using the HRM version 2.0.1 software (Applied Biosystems), followed by clustering analysis as described by Reja et al. [14].

2.7. SNP genotyping

2.7.1. TaqMan genotyping

Specific TaqMan genotyping assays (ID: C__25766207_10; Thermo Fisher Scientific, Waltham, MA, USA) and the 7500 Fast RT-PCR System (Applied Biosystems) was used to determine the rs3796704 (*ENAM*) allele distribution in the analysed population. TaqMan genotyping was performed using 8 ng of DNA, 5 µL of the TaqMan Universal Master Mix II (Thermo Fisher Scientific, Waltham, MA, USA), 0.25 µL TaqMan assay and 3.75 µL of nuclease-free water. The PCR temperature protocol comprised polymerase activation 95 °C for 10 min followed by 40 cycles of denaturation 95 °C for 15 s, annealing/extension at 60 °C for 1 min, followed by an endpoint plate read and allele discrimination analysis using the 7500 software version 2.0.6 (Applied Biosystems).

2.7.2. HRM analysis

Selected SNPs in the *ENAM* interacting genes (Table 1) were detected using HRM analysis performed on the 7500 Fast RT-PCR System (Applied Biosystems, USA) using Type-IT HRM master mix (Qiagen, Germany). Each HRM reaction consisted of 20 ng of DNA, 10 µL of the Type-IT HRM Master Mix, 0.4 µL forward primer, 0.4 µL reverse primer and 5.2 µL of nuclease-free water. The PCR protocol comprised polymerase activation at 95 °C for 5 min, followed by 50 cycles of denaturation at 95 °C for 10 s, annealing at 60 °C for 25 s, and extension at 72 °C for 15 s. Melting curves were recorded between 70 °C and 98 °C with a 1% temperature ramp, and were aligned and normalised using the HRM version 2.0.1 software (Applied Biosystems), followed by clustering analysis as described by Reja et al. [14].

2.7.3. Sanger sequencing

Two samples with each variant detected by HRM approach were chosen for further sequence analysis of selected SNPs in the *ENAM* interacting genes. The selected HRM samples were diluted in the ratio 1:4 with nuclease-free water. Excess/unused primers and nucleotides from the HRM reaction were hydrolysed using ExoSAP-IT™ (Affymetrix, Santa Clara, CA, USA). Each reaction consisted of 2.5 µL of HRM amplicon and 1 µL of ExoSAP-IT™, followed by incubation at 37 °C for 15 min to degrade remaining primers and nucleotides. ExoSAP-IT™ was inactivated by incubation at 80 °C for 15 min. 16.2 µL of sequencing master mix BigDye® Terminator v3.1 (Applied Biosystems) and 0.3 µL forward or reverse primer was added to each sample. GeneAmp® PCR System 9700 (Applied Biosystems) was used for sequence reaction with polymerase activation at 96 °C for 1 min, followed by 25 cycles of denaturation at 96 °C for 10 s, annealing at 50 °C for 5 s, and extension at 60 °C for 4 min. After the sequence reaction unincorporated ddNTPs were removed with 3 M sodium acetate and ethanol precipitation. Precipitated and dried products of the sequence reaction were later on dissolved in formamide. The capillary electrophoresis was performed on the ABI 3500 Genetic Analyser (Applied Biosystems) and sequences were analysed with the Sequencing Analysis Software v5.4. (Applied Biosystems).

2.8. Statistical analysis

Based on the data presented here, we performed a statistical analysis which will be published in the associated research article [15].

No additional statistical analysis was performed.

Ethics statement

Written informed consent from parents or legal guardians to allow their children to participate in all procedures associated with the study was obtained.

The study was approved by the Slovenian National Medical Ethics Committee (65/05/14).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

CRediT authorship contribution statement

Luka Hočevar: Investigation, Writing - original draft, Visualization, Writing - review & editing. **Jernej Kovač:** Investigation, Formal analysis, Resources, Writing - review & editing. **Katarina Trebušak Podkrajšek:** Investigation, Formal analysis, Resources, Writing - review & editing. **Saba Battelino:** Conceptualization, Writing - review & editing. **Alenka Pavlič:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Supervision.

Acknowledgements

We are grateful to all the participants and their parents/guardians. We also thank the anaesthesia technicians for blood collection, and Ms. Jurka Ferran for support in the laboratory.

This work was supported by the Slovenian Research Agency (Grant no: 20150115, 2015).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2020.106224.

References

[1] A. Ghanim, M. Elfrink, K. Weerheijm, R. Mariño, D. Manton, A practical method for use in epidemiological studies on enamel hypomineralisation, Eur. Arch. Paediatr. Dent. 16 (2015) 235–246, doi:10.1007/s40368-015-0178-8.

[2] String, Protein-Protein Interaction Networks Functional Enrichment Analysis. https://version-10-5.string-db.org/, 2018 (accessed 19 September 2018).

[3] B. Snel, G. Lehmann, P. Bork, M.A. Huynen, STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene, Nucl. Acids Res. 28 (2000) 3442–3444, doi:10.1093/nar/28.18.3442.

[4] D. Szklarczyk, A.L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N.T. Doncheva, J.H. Morris, P. Bork, L.J. Jensen, C.V. Mering, STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, Nucl. Acids Res. 47 (2019) D607–D613, doi:10.1093/nar/gky1131.

[5] CADD, Combined Annotation Dependent Depletion. https://cadd.gs.washington.edu, 2018 (accessed 19 September 2018).

[6] K.L. Weerheijm, M. Duggal, I. Mejàre, L. Papagiannoulis, G. Koch, L.C. Martens, A.L. Hallonsten, Judgement criteria for molar incisor hypomineralisation (MIH) in epidemiologic studies: a summary of the European meeting on MIH held in Athens, 2003, Eur. J. Paediatr. Dent. 4 (2003) 110–113.
[7] F. Jeremias, M. Koruyucu, E.C. Küchler, M. Bayram, E.B. Tuna, K. Deeley, R.A. Pierri, J.F. Souza, C.M. Fragelli, M.A. Paschoal, K. Gencay, F. Seymen, R.M. Caminaga, L. dos Santos-Pinto, A.R. Vieira, Genes expressed in dental enamel development are associated with molar-incisor hypomineralization, Arch. Oral. Biol. 58 (2013) 1434–1442, doi:10.1016/j.archoralbio.2013.05.005.
[8] M. Kircher, D.M. Witten, P. Jain, B.J. O'Roak, G.M. Cooper, J. Shendure, A general framework for estimating the relative pathogenicity of human genetic variants, Nat. Genet. 46 (2014) 310–315, doi:10.1038/ng.2892.
[9] P. Rentzsch, D. Witten, G.M. Cooper, J. Shendure, M. Kircher, CADD: predicting the deleteriousness of variants throughout the human genome, Nucl. Acids Res. 47 (2019) D886–D894, doi:10.1093/nar/gky1016.
[10] ENSEMBL database, Ensembl Genome Browser. http://grch37.ensembl.org/Homo_sapiens/Info/Index, 2018 (accessed 19 September 2018).
[11] Primer3, http://bioinfo.ut.ee/primer3-0.4.0/, 2018 (accessed 19 September 2018).
[12] SNP check. https://genetools.org/SNPCheck/snpcheck.htm, 2018 (accessed 19 September 2018).
[13] Primer-BLAST. Primer Basic Local Alignment Search Tool Interface. https://www.ncbi.nlm.nih.gov/tools/primer-blast/, 2018 (Accessed 19 September 2018).
[14] V. Reja, A. Kwok, G. Stone, L. Yang, A. Missel, C. Menzel, B. Bassam, ScreenClust: advanced statistical software for supervised and unsupervised high resolution melting (HRM) analysis, Methods 50 (2010) 10–14, doi:10.1016/j.ymeth.2010.02.006.
[15] L. Hočevar, J. Kovač, K. Trebušak Podkrajšek, S. Battelino, A. Pavlič, The possible influence of genetic aetiological factors on molar–incisor hypomineralisation, Arch. Oral Biol. 118 (2020) 104848, , doi:10.1016/j.archoralbio.2020.104848.