



HAL
open science

Neanderthal and Denisova tooth protein variants in present-day humans

Clément Zanolli, Mathilde Hourset, Rémi Esclassan, Catherine Mollereau

► **To cite this version:**

Clément Zanolli, Mathilde Hourset, Rémi Esclassan, Catherine Mollereau. Neanderthal and Denisova tooth protein variants in present-day humans. PLoS ONE, 2017, 12 (9), pp.e0183802. 10.1371/journal.pone.0183802 . hal-02296655

HAL Id: hal-02296655

<https://hal.science/hal-02296655>

Submitted on 26 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

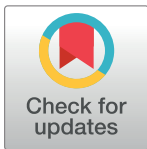
RESEARCH ARTICLE

Neanderthal and Denisova tooth protein variants in present-day humans

Clément Zanolli^{1*}, Mathilde Hourset^{1,2}, Rémi Esclassan^{1,2}, Catherine Mollereau^{1*}

1 Laboratoire d'Anthropologie Moléculaire et Imagerie de Synthèse (AMIS), UMR5288 CNRS - Université de Toulouse, Toulouse, France, **2** Faculté de chirurgie dentaire, Université de Toulouse, Toulouse, France

* clement.zanolli@gmail.com (CZ); catherine.mollereau-manaute@ipbs.fr (CM)



Abstract

Environment parameters, diet and genetic factors interact to shape tooth morphostructure. In the human lineage, archaic and modern hominins show differences in dental traits, including enamel thickness, but variability also exists among living populations. Several polymorphisms, in particular in the non-collagenous extracellular matrix proteins of the tooth hard tissues, like amelogenin, are involved in dental structure variation and defects and may be associated with dental disorders or susceptibility to caries. To gain insights into the relationships between tooth protein polymorphisms and dental structural morphology and defects, we searched for non-synonymous polymorphisms in tooth proteins from Neanderthal and Denisova hominins. The objective was to identify archaic-specific missense variants that may explain the dental morphostructural variability between extinct and modern humans, and to explore their putative impact on present-day dental phenotypes. Thirteen non-collagenous extracellular matrix proteins specific to hard dental tissues have been selected, searched in the publicly available sequence databases of Neanderthal and Denisova individuals and compared with modern human genome data. A total of 16 non-synonymous polymorphisms were identified in 6 proteins (ameloblastin, amelotin, cementum protein 1, dentin matrix acidic phosphoprotein 1, amelogenin and matrix Gla protein). Most of them are encoded by dentin and enamel genes located on chromosome 4, previously reported to show signs of archaic introgression within Africa. Among the variants shared with modern humans, two are ancestral (common with apes) and one is the derived amelogenin major variant, T648I (rs7671281), associated with a thinner enamel and specific to the *Homo* lineage. All the others are specific to Neanderthals and Denisova, and are found at a very low frequency in modern Africans or East and South Asians, suggesting that they may be related to particular dental traits or disease susceptibility in these populations. This modern regional distribution of archaic dental polymorphisms may reflect persistence of archaic variants in some populations and may contribute in part to the geographic dental variations described in modern humans.

OPEN ACCESS

Citation: Zanolli C, Hourset M, Esclassan R, Mollereau C (2017) Neanderthal and Denisova tooth protein variants in present-day humans. PLoS ONE 12(9): e0183802. <https://doi.org/10.1371/journal.pone.0183802>

Editor: Alistair Robert Evans, Monash University, AUSTRALIA

Received: April 5, 2017

Accepted: August 13, 2017

Published: September 13, 2017

Copyright: © 2017 Zanolli et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. They were collected from public databases shared with the scientific community.

Funding: The authors received funding from the CNRS and IFRO.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: AI, amelogenesis imperfecta; AMBN, ameloblastin; AMELX, chromosome X encoded amelogenin; AMELY, chromosome Y

encoded amelogenin; AMH, anatomically modern humans; AMTN, amelotin; CEMP1, cementoblastoma-derived protein/cementum protein 1; DMP1, dentin matrix acidic phosphoprotein 1; DSPP, dentin sialophosphoprotein; EDA, ectodysplasin-A; ENAM, enamelin; KLK4, kallikrein-4/enamel matrix serine proteinase 1; MGP, matrix Gla protein; MMP20, matrix metalloproteinase-20/enamel metalloproteinase; TUFT1, tuftelin.

Introduction

Dental morphology shows variability between extant and extinct human groups, but also among modern humans according to a regional distribution [1–3]. Tooth characteristics are shaped by different factors, including of course genetic factors governing dentine and enamel tissue formation [4–8] (<http://bite-it.helsinki.fi>). But life history parameters, environmental conditions and diets should also be taken into account when studying dental morphostructural evolutionary trends [9, 10]. In the present study, we aimed to determine whether archaic-specific missense variants in tooth proteins from Neanderthal and Denisovan individuals could explain past and present dental morphostructural variability and/or disease susceptibility.

Teeth are composed of three hard tissues (enamel, dentine and cementum) surrounding the pulp and expressing specific non-collagenous proteins including extracellular matrix proteins and proteases involved in the control of mineralization process and crystal deposition during tooth development [11, 12]. Enamel is the hardest external tissue dedicated to mastication. It is formed before tooth eruption by secretory ameloblasts through a two-step process. During the secretion phase of amelogenesis, unique matrix proteins are released (mainly amelogenin, ameloblastin, enamelin, tuftelin and amelotin) and participate to the initiation and growth of hydroxyapatite crystals [11, 13]. During the maturation phase, the matrix proteins are degraded by the specific proteases matrix metalloproteinase 20 and kallikrein 4. Dentin, the “shock absorber” between enamel and pulp, is produced by odontoblast cells lying at the periphery of the pulp where they also play a role in defense and sensory transmission [11]. The dentin organic matrix is essentially composed of collagens but it also contains some proteins controlling dentin mineralization such as the dentin matrix acidic phosphoprotein 1 and the dentin sialophosphoprotein-derived proteins, dentin sialoprotein, dentin glycoprotein and dentin phosphoprotein involved in the nucleation and growth of mineral crystals. The cementum is a thin mineralized layer produced by cementoblasts. It covers the tooth root and maintains its attachment to the alveolar bone. It also acts as a microbial barrier [11, 14]. The composition of the cementum organic extracellular matrix shows similarities with that of the alveolar bone. It is rich in growth factors and glycoproteins that participate in periodontal tissue repair and regeneration after damage. The cementum protein 1 is specific to cementum. It plays a role in the early steps of mineralization and in the local regulation of cell differentiation and proliferation [14]. The bone matrix Gla protein is not specific to cementum but is expressed in the cementum mineralization front where it exerts an inhibitory control on calcification essential for proper mineralization.

The selective pressure and adaptation to specific diets is based on a balance between sharp-cutting efficiency, tooth resistance and protection, and likely played a role in the selection of enamel thickness in primates [15–17]. Fossil and extant hominoids represent an extreme example. To the exception of a few hard-feeding Miocene and Pliocene hominids like *Ouranopithecus* [18] and australopiths [19–21], the genus *Homo* exhibits thicker enamel, with little variation through the course of its evolution during the Pleistocene [20, 22, 23]. However, some human groups like Neanderthals tend to show a thinner enamel [22, 24]. Thinner enamel is also observed for the c. 1 million-years-old Eritrean *H. erectus/ergaster* dental sample from Buia and Mulhuli Amo [25].

Many other traits distinguish Neanderthal and modern humans [1, 22, 26, 27]. Neanderthals exhibit a particular anterior dental morphology characterized by a higher degree of robustness and labial convexity, and a higher frequency of shovel-shaped incisors with a marked tuberculum dentale (in the upper incisors), compared to modern humans or African hominin fossils [1, 28]. The Neanderthal post-canine teeth also display a number of typical traits, including a high frequency of accessory features, the lingual displacement of the

hypocone, a well-developed mid-trigonid crest (also expressed at the enamel-dentin junction level), centrally set protoconid and entoconid dentin horns, taurodontic roots and enlarged pulp cavity [1, 27, 29]. However, Neanderthals share with modern humans a tendency to reduce their posterior dentition [28] and a similar molar root growth timing [27], even if the tooth crowns tend to show faster dental maturation in Neanderthals [30]. But the frequency of accessory traits and their level of expression, such as for the middle trigonid crest or the centrally shifted dentin horns, are generally much lower in modern humans, at both the outer enamel surface and at the enamel-dentin junction level [27, 29, 31]. Regarding the Denisovan hominins, another extinct human group recently identified by ancient DNA sequencing, dental morphological variability is less documented as only 2 molars have been described, Denisova 4 and 8 [32, 33]. These teeth exhibit distinct primitive morphological characteristics, including the presence of numerous accessory cusps. They are larger than the molars of Neanderthals, modern humans and Asian archaic hominins such as *Homo erectus*, but share with the later a trapezoidal shape [32, 33]. The enamel-dentin junction of the maxillary molar Denisova 8 exhibits a unique occlusal topography, with a large hypocone, a well-developed metacornule and hypercomplex expression of accessory crests and cusps [33]. However, no information on enamel thickness is available to this date on the Denisovan hominin teeth.

The shape of the outer enamel surface is determined by the enamel-dentin junction through growth of the enamel cap [21, 34]. Polymorphisms in genes or proteins involved in the formation of this tooth region are therefore expected to participate in tooth variability. A number of mutations in the proteins involved in the mineralization (AMEL, AMBN, AMTN, ENAM) or the maturation processes (DSPP, KLK4, MMP20) of enamel and dentin are indeed associated with particular enamel phenotypes or tooth defects, or create susceptibility to dental disorders such as amelogenesis imperfecta (AI), dentinogenesis imperfecta (DI)/dysplasia (DD) or caries [12, 16, 35–42] (Table 1). Recent genetic evidence suggests that enamelin, in tandem with regulation of the processing of the enamel matrix proteins by MMP20, is linked to the evolution of enamel thickness in humans [16]. Conversely, a non-synonymous polymorphism in the enamelin protein, T648I (rs7671281), is associated with a thinner enamel in living humans [15]. In addition, a Neanderthal variant in the post-transcriptional repressor miR-1304 leading to a reduced expression of enamelin and amelotin genes could explain the variation in enamel thickness between Neanderthals and modern humans [26]. Moreover, external tooth morphology may be also affected by a single protein mutation. This is the case of the Asian-specific non-synonymous polymorphism (V370A, rs3827760) in the ectodysplasin A receptor gene (EDAR) which is associated with the expression of shoveling and double-shoveling in upper first incisors, a particular trait of extant Asian populations [43]. This variant, in addition to T648I (rs7671281) in the enamelin protein, emphasizes the role a single amino acid change may play in tooth morphostructure.

In light of these data, the present study aimed to investigate whether tooth protein variants in archaic hominins could provide valuable molecular information to help understand past and present human dental morphostructural variability. To do this, the Neanderthal and Denisovan sequences of the main non-collagenous extracellular matrix proteins of tooth hard tissues (listed in Table 1) were retrieved from publicly available databases [44–46] and were compared with modern human proteins. The detected variants were analyzed for their ancestry, their putative functional impact, and their frequency and regional distribution in living populations, based on informations collected from online databases and the literature. Moreover, finding archaic polymorphisms in present-day humans may enable us to trace back the legacy from archaic admixture with anatomically modern humans (AMH) some 50 000 years ago [47], bringing molecular clues to explain particular dental morphologies or disorders in living populations.

Table 1. Main non-collagenous extracellular matrix proteins present in tooth hard tissues.

Gene	Protein		Function	Phenotype linked to polymorphism in humans
Name	Name	Uniprot Access.N ^o		
AMBN	Ameloblastin	Q9NP70	Structural constituent of enamel, involved in mineralization	AI [40]
AMELX	Amelogenin X	Q99217	Structural constituent of enamel, involved in mineralization	Hypoplasia, AI [12, 42]
AMELY	Amelogenin Y	Q99218	Structural constituent of enamel, involved in mineralization	
AMTN	Amelotin	Q6UX39	Structural constituent of enamel, involved in mineralization	Hypomineralization, AI [41]
CEMP1	Cementum protein 1	Q6PRD7	Structural constituent of cementum	
DMP1	Dentin matrix acidic phosphoprotein 1	Q13316	Structural constituent of dentin	
DSPP	Dentin sialophosphoprotein	Q9NZW4	Structural constituent of dentin, involved in mineralization	DI, DD [12, 38]
ENAM	Enamelin	Q9NRM1	Structural constituent of enamel, involved in mineralization	Enamel thickness [15] Hypoplasia, AI [36, 37, 42]
KLK4	Kallikrein-4 (Enamel matrix serine proteinase 1)	Q9Y5K2	Involved in the maturation of enamel matrix proteins	Hypoplasia, AI [12]
MMP20	Matrix metalloproteinase-20 (Enamelysin)	O60882	Involved in the maturation of tooth matrix proteins	Hypoplasia AI [12]
MGP	Matrix Gla protein	P08493	Structural constituent of bone and tooth, negative regulator of mineralization	
ODAM	Odontogenic ameloblast-associated protein	A1E959	Structural constituent of enamel	
TUFT1	Tuftelin	Q9NNX1	Structural constituent of enamel	

AI: amelogenesis imperfecta, DD: dentinogenesis dysplasia, DI: dentinogenesis imperfecta

<https://doi.org/10.1371/journal.pone.0183802.t001>

Methods

The coding sequences of 13 non-collagenous proteins characteristic of tooth tissues (listed in Table 1) were extracted from the publicly available exome database [44] that includes three Neanderthal specimens, the Altaï pedal phalanx (Altaï Mountain Cave, Siberia) and the bone fragments Vi33.15 (Vindija cave, Croatia) and SD1253 (El Sidron cave, Spain). The sequences were aligned to the reference *Homo sapiens* protein sequences in Uniprot (<http://www.uniprot.org/>). In addition, the tooth protein variants in Denisova [45] and in the incomplete reads of other Neanderthal individuals were searched in the Neanderthal Genome Analysis Consortium Tracks on the UCSC Genome Browser website (<https://genome-euro.ucsc.edu/>; GRCh37/hg19 Assembly) [48]. Among the 13 proteins, amelogenins, particularly the Y-encoded form (AMELY), could not be correctly analyzed because of a poor sequence coverage. Likewise, the sequence of the cementum protein 1 was only available in the Denisova genome. For each variant found, the amino acid change was compared with the chimpanzee (*Pan troglodytes*) and gorilla (*Gorilla gorilla*) proteins in Uniprot in order to assess ancestry. The frequency and the putative phenotype of the protein polymorphism in present-day people was searched in the public Ensembl project database (<http://www.ensembl.org/>; Ensembl release 84-March 2016) [49]. The geographical distribution maps of the dental protein variants were generated by querying each polymorphism in the online Geography of Genetic Variants

Browser (<http://popgen.uchicago.edu/ggv/>) that compiles Single Nucleotide Variant information from public genomic datasets such as the 1000 Genome Project Consortium [50].

In the present study: the term archaic refers to the extinct hominins Neanderthal and Denisova, the term ancestral indicates that the variant is shared with chimpanzee and gorilla; conversely, the term derived indicates that the variant is not ancestral and differs from apes.

Results

The alignment to the reference human genome of the protein sequences of the 13 tooth non-collagenous extracellular matrix proteins of the archaic hominins (Neanderthals and Denisova) provided a total of 16 non-synonymous polymorphisms in 6 proteins (Table 2). They are essentially found in Denisova and the Neanderthal individuals Altai and Vi33.15, owing to their highest sequence coverage. The alignment with the sequences of chimpanzee and gorilla indicates that only two amino acid changes, K55E (rs13331643) in CEMP1 and T127A (rs4236) in MGP are common with apes (Table 2). These variants are also present in modern humans (7% and 39% frequency, respectively) and likely corresponds to ancestral variants. All the other archaic amino acid changes (Table 2) correspond to derived variants found only at low frequencies (< 1%) in living populations, except for the enamel substitution T648I (rs7671281) which is a major polymorphism today (90%). Conversely, the major alleles in present-day humans are shared with apes, except K55 in CEMP1 that could be considered as modern-derived, and I648 in enamel, which is common to modern and archaic humans and could be considered as *Homo*-specific. Archaic hominins thus appear to have expressed particular teeth protein variants with likely a high frequency as some of them are shared by several archaic individuals although unrelated in time and geographic location (Table 2). A detailed description of these non-synonymous polymorphisms and their potential functional impact are discussed below.

Geographical distribution and frequency of the polymorphisms shared by archaic hominins

The 3 non-synonymous polymorphisms common to Neanderthals and Denisova are frequently found in people today (Table 2 and Fig 1). The amino acid substitution T648I (rs7671281) in ENAM corresponds to the major derived variant, nearly fixed out of Africa. T127A (rs4236) in MGP is ancestral and commonly found today, however with the smallest frequency in East Asia. A8V (rs142330429) in MGP is found at a frequency of 3–4% in Asian population only, although a few cases are found in Finland (Fig 1). Interestingly, the two polymorphisms in MGP are detected in Denisovan and Neanderthal individuals arising from nearly all the geographical sites (Russia, Croatia and Spain) indicating a high frequency in archaic hominins. This could be expected for T127A (rs4236) which is an ancestral polymorphism, and suggests that the variant A8V (rs142330429) was probably present in a common ancestor before the separation of Neanderthal and Denisovan lineages that occurred around 381,000–473,000 years ago [51]. Persistence of the A8V polymorphism in East and South Asia may suggest a local segregation in this region.

Geographical distribution and frequency of the polymorphisms specific to Neanderthal or Denisova

In contrast to the non-synonymous polymorphisms shared by archaic hominins, those specific to either Neanderthals (Fig 2) or Denisova (Fig 3) correspond, for most of them, to non-ancestral variants that are currently present at low or very low frequency in living humans (Table 2).

Table 2. Non-synonymous polymorphisms present in tooth proteins of Neanderthal and Denisova hominins.

Variants identified in archaic hominins	Gorilla	Chimp	Modern humans (global worldwide frequency)	Neanderthals						El Sidron (Spain)	Denisova
				Altaï (Russia)	Vindija (Croatia)				Mezmaiskaya (Russia) Mez1		
					Vi33.15	Vi33.16	Vi33.25	Vi33.26			
Ameloblastin (AMBN)											
G78S (rs143795139) derived	G	G	S (0.0001)	G	S	G	G	G		G	G
M273V (rs564905233) derived	M	M	V (0.001)	M	M	M					V
Amelotin (AMTN)											
R171H (rs117133443) derived	R	R	H (0.0001)	R	R	R	R				H
A200V (rs376682442) derived	T	A	V (0.005)	A	A			A		A	V
Cementum protein 1 (CEMP1)											
R15Q (rs370735504) derived		R	Q (0.0001)					R	R		Q
K55E (rs13331643) ancestral		E	E (0.07)								E
R80H (rs79379654) derived		R	H (0.001)				R				H
Dentin matrix acidic phosphoprotein 1 (DMP1)											
R173Q (rs747698893) derived	R	R	Q (0.0001)	R	R		R				Q
N483T (rs574215585) derived	N	N	T (0.005)	T	T			T			N
Enamelin (ENAM)											
G389S (rs74511578) derived	G	G	S (0.01)	S	S		S	S			G
T415I (rs779308123) derived	T	T	I (0.0001)	T	T		T	T	I	T	T
R429H (rs375758699) derived	R	R	H (0.0005)	R	R					R	H
T648I (rs7671281) derived	T	T	I (0.82)	I	I						I
T951M (rs561402393) derived	T	T	M (0.0001)	T	M		M	M		T	T
Matrix Gla protein (MGP)											
A8V (rs142330429) derived	A	A	V (0.01)	V	V		V	V		V	V
T127A (rs4236) ancestral		A	A (0.39)	A	A	A			A	A	A

Bold characters indicate the aminoacid and position corresponding to the archaic hominin variant. Empty cell indicates no sequence read available at this position. Yellow: common to archaic hominins; Dark green: specific to Neanderthals; Light green: specific to the Vindija cave; Dark blue: specific to Denisova; Light blue: at least present in Denisova.

<https://doi.org/10.1371/journal.pone.0183802.t002>

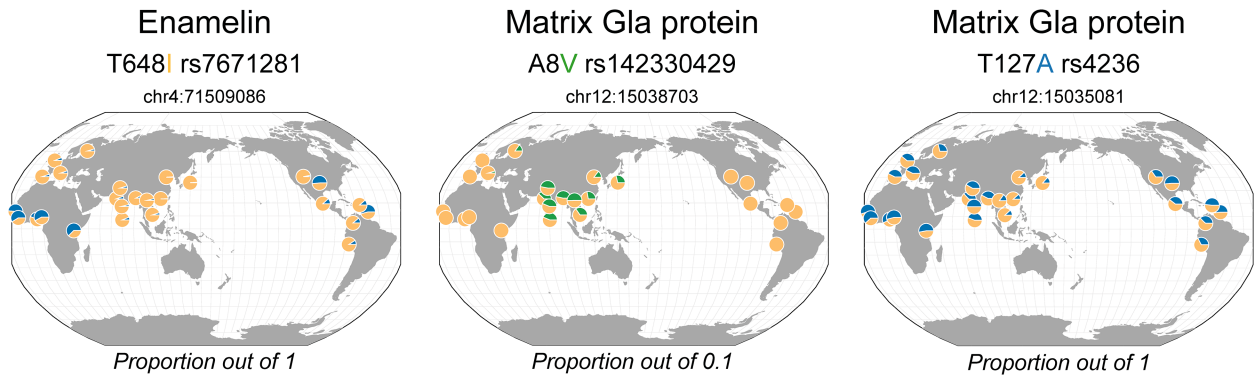


Fig 1. Geographical distribution of the dental protein variants common to archaic hominins in living people. The colored amino acid corresponds to the archaic variant. Yellow corresponds to the major allele; blue and green correspond to minor alleles with a proportion out of 1 and 0.1, respectively. The maps were generated in the Geography of Genetic Variants Browser (<http://popgen.uchicago.edu/ggv/>) [50].

<https://doi.org/10.1371/journal.pone.0183802.g001>

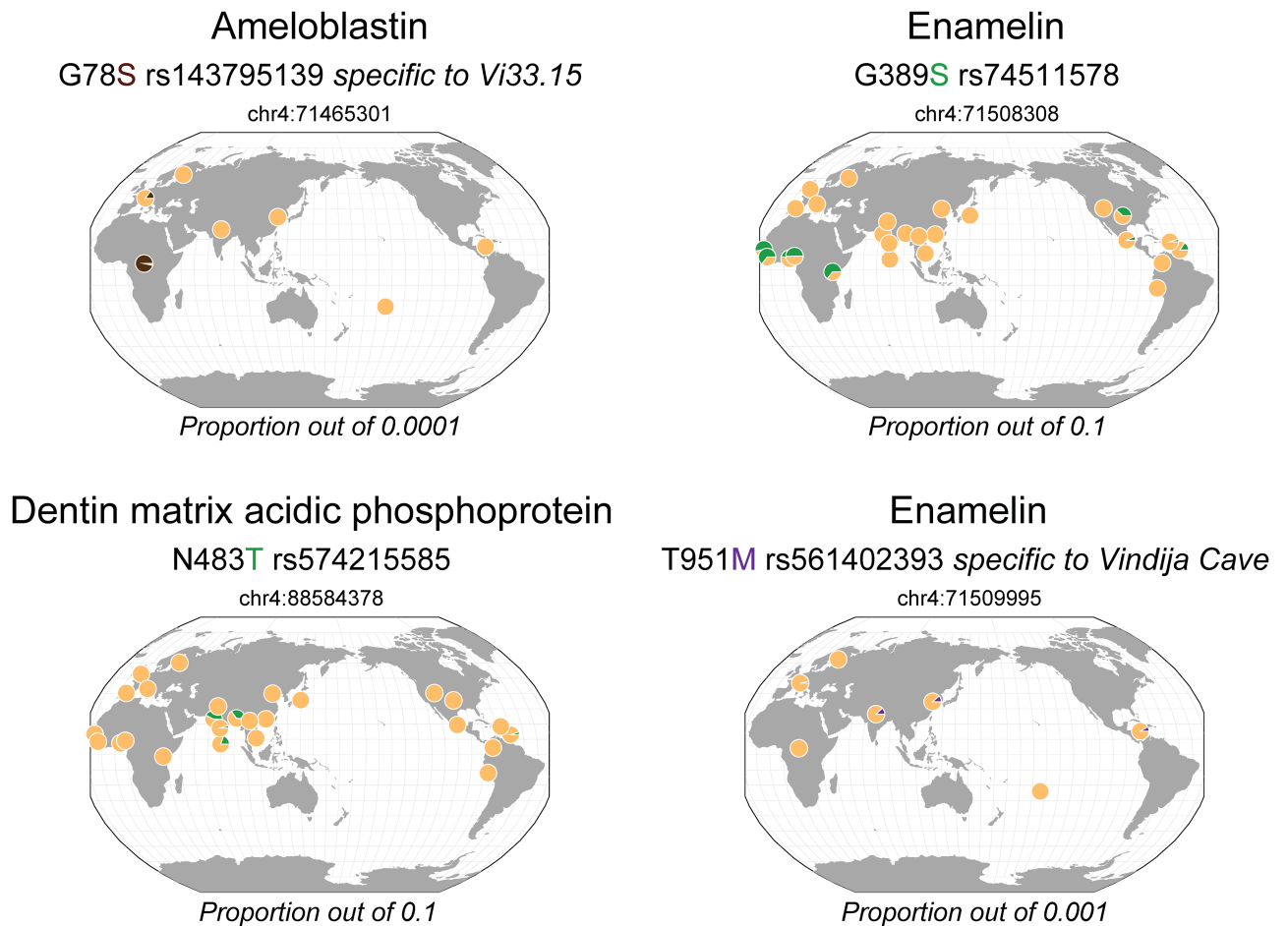


Fig 2. Geographical distribution of the dental protein variants specific to Neanderthals in living people. The colored amino acid corresponds to the archaic variant. Yellow corresponds to the major allele, green, purple and brown correspond to minor alleles with a proportion out of 0.1, 0.001 and 0.0001, respectively. The maps were generated in the Geography of Genetic Variants Browser (<http://popgen.uchicago.edu/ggv/>) [50].

<https://doi.org/10.1371/journal.pone.0183802.g002>

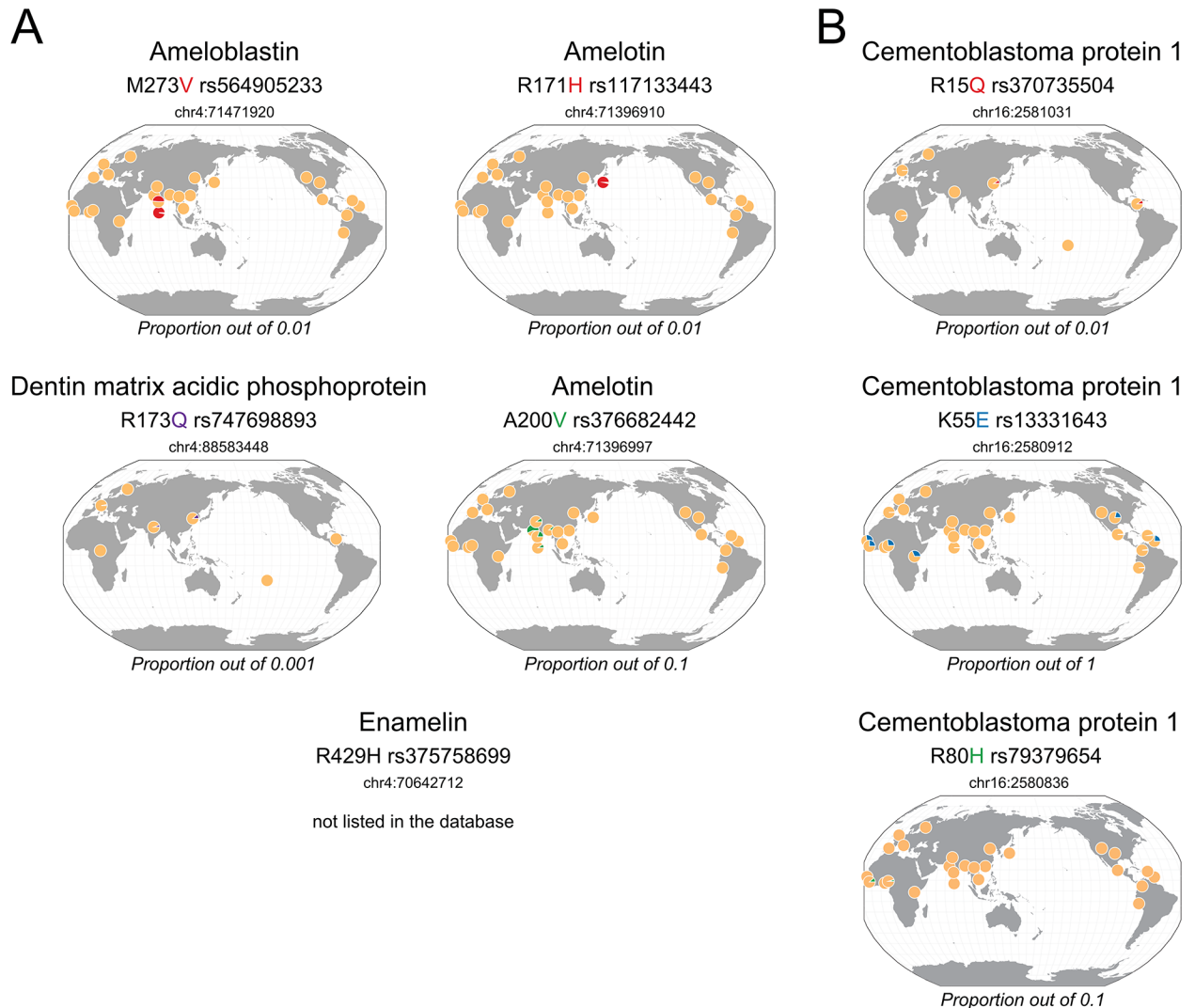


Fig 3. Geographical distribution of dental protein variants A: specific to Denisova, B: at least present in Denisova in living people. The colored amino acid correspond to the archaic variant. Yellow corresponds to the major allele, blue, green, red and purple corresponds to minor alleles with a proportion out of 1, 0.1, 0.01 and 0.001, respectively. The maps were generated in the Geography of Genetic Variants Browser (<http://popgen.uchicago.edu/ggv/>) [50].

<https://doi.org/10.1371/journal.pone.0183802.g003>

These polymorphisms are, however, mostly found in Africa or in East and South Asia populations (S1 Table).

In Neanderthals, the individuals from the Altai and Vindija caves express the derived amino acid changes N483T (rs574215585) in DMP1 and G389S (rs74511578) in ENAM. Nowadays, the respective frequency of these variants is 2% in South Asia and 5% in Africa and America (Fig 2). The variant G78S (rs143795139) in AMBN is specific to the Vi33.15 individual and is observed today at very low ($\leq 0.01\%$) frequency in Africa and South Europe. In addition, another substitution in ENAM, T951M (rs561402393), is specific to the individuals from the Vindija cave and corresponds to a rare variant ($< 0.1\%$) present today in South and East Asia as well as South America.

The tooth proteins of the Denisova individual contain a large number of specific polymorphisms: M273V (rs564905233) in AMBN; R171H (rs117133443) and A200V (rs376682442) in

AMTN; R173Q (rs747698893) in DMP1 that are essentially found today in the Asian population at a low (1–5%) or very low (0.01%) frequency. In addition, three variants in CEMP1, R15Q (rs370735504), K55E (rs13331643) and R80H (rs79379654), are also detected in Denisova, but cannot definitively be attributed to this lineage in the absence of protein coverage in the Neanderthal individuals (Fig 3B). K55E is a common variant in Africa (26% frequency, Fig 3 and S1 Table).

Archaic tooth protein variants: Description and discussion

Enamelin. It is interesting to note that the majority of the non-synonymous polymorphisms identified in archaic hominins are found in enamel (ENAM), the protein playing an essential role in the proper formation of enamel and involved in enamel thickness [15]. It has been proposed that the ENAM gene was under positive selection related to diet change during the primate lineage evolution, and some enamel variants were shown to influence enamel thickness [15–17]. In particular, the derived ENAM amino acid substitution T648I (rs7671281), which is common in sub-Saharan Africa and nearly fixed out-of Africa (Table 2 and Fig 1), is associated with a thinner enamel [15] and may explain the difference in dental size between populations with African and non-African origins [2, 3, 52]. Interestingly, this polymorphism is present in both Neanderthal and Denisova hominins and not apes, suggesting that it was carried by a more ancient common ancestor, consistent with positive selection in the *Homo* lineage [17]. This variant may have contributed to the continuous decrease in enamel thickness from Plio-Pleistocene hominins (including *Australopithecus* and *Paranthropus*) to *Homo* [53].

However, the T648I polymorphism by itself cannot fully explain the relatively thinner enamel of Neanderthals as this variant is shared with modern humans that exhibit a thicker enamel [22, 24] and with Denisovans that harbor tooth characteristics more reminiscent to early *Homo* [32, 33]. It is therefore possible that the other polymorphisms specific to each hominin lineage may explain some of these differences. In particular, most of the Neanderthal individuals show accumulation of specific amino acid changes in ENAM (G389S, T951M) that could have contributed, in addition to another Neanderthal specific substitution in DMP1 (N483T, Table 2) and to the miR-1304 polymorphism described by [26], to the singularity of Neanderthal teeth characterized by an apparent thin enamel and particular dentin/enamel proportions [24, 29]. A difference in tooth size has been proposed to explain variation in enamel thickness. Modern humans have smaller molars than Neanderthals but the coronal dentine is disproportionately reduced compared to the enamel cap, resulting in relatively higher enamel thickness values [22, 24]. How proteins may affect the relative proportions of dentin and enamel in relation to crown size is yet unknown. Investigating the influence of the ENAM and DMP1 polymorphisms identified here on dental structure, especially in the Asian population where they are more frequent, could be a mean to get more insights.

Furthermore, the T648I variant in ENAM has been proposed to offer protection against caries since the reverse variant I648T, corresponding to the ancestral and today minor allele, has been associated, in combination with another ENAM polymorphism R736Q, with increased risk of caries in a French children cohort [35]. The advantageous T648I substitution nearly fixed today may have been inherited from the archaic hominins, that show rare sign of caries [54, 55] and low level of defective enamel formation [37]. In addition to susceptibility to caries, mutations in the ENAM gene have been described to cause enamel defects such as hypoplasia (thinner and more irregular enamel) and hypomineralization linked to *Amelogenesis imperfecta* (AI) disorders [37]. Local hypoplastic enamel defect occurs from <1% to >56%

of living humans depending on populations, and was rare in Neanderthals [37] and earlier human groups [56, 57].

The 5 non-synonymous ENAM polymorphisms found in archaic hominins (Table 2) are all located in the more variable region encoded by exon 10 of the protein and do not correspond to any described SNP in AI disease [58]. Four of these mutations lead to the addition or the removal of a putative phosphorylation site in the protein. It is interesting to note that one reported case of AI corresponds to the suppression of a phosphorylation site on the S216 residue of ENAM [59], suggesting that the variants identified in the archaic hominins may have a functional impact.

Dentin matrix acidic phosphoprotein 1. DMP1 is expressed in odontoblasts and osteoblasts and regulates cell differentiation and biomineralization of dentin and bones. The protein acts as a transcriptional factor regulating the expression of bone and dental genes, and also participates in phosphate homeostasis [11, 60]. Deletion or missense mutation of the protein lead to cartilage malformation, defect in the pre-dentin to dentin maturation, hypomineralization and hypophosphatemia [60]. The Neanderthal (N483T) or Denisova (R173Q) variants have not been shown involved in a particular phenotype or disease. They are located in the C-terminal of DMP1 that is enzymatically-processed to a 57 kDa fragment promoting hydroxyapatite nucleation and containing nuclear localization signals. Interestingly, N483T is located in a highly phosphorylated domain of the protein [60]. DMP1 phosphorylation is important for the control of dentin and bone matrix mineralization. Therefore, the threonine residue introduced by the mutation could potentially create an additional phosphorylation site susceptible to have functional consequence, such as a higher collagen matrix mineralization.

Ameloblastin. After amelogenin, ameloblastin (AMBN) is the second most abundant enamel matrix protein. The protein interacts with amelogenin and participates in enamel formation by a not yet elucidated mechanism [11, 13, 40]. Nevertheless, AMBN plays an essential role in amelogenesis as demonstrated by the severe hypoplasia of knock-out mice [61]. Moreover, a deletion in the AMBN gene has recently been involved in a case of AI in a Costa Rica family [40]. Apart from this deletion, the functional consequence of AMBN polymorphisms are poorly documented. The two variants identified in Neanderthal Vi3315 (G78S) and in Denisova (M273V) individuals are not located into any functional or conserved domains of the protein [62].

Amelotin. This protein has been recently identified as another, less abundant enamel matrix protein expressed during the maturation stage of amelogenesis and was proposed to promote hydroxyapatite precipitation [13, 41]. As for AMBN, amelotin knock-out mice show enamel defects and deletion of exons 3–6 in the AMTN gene has been associated with AI in a Costa Rica family [41]. Two AMTN amino acid changes (R171H, A200V) were detected in the Denisova individual only and may form a haplotype characteristic of this lineage. The variants are not located in protein domains conserved during mammalian evolution or harboring biological activity [63].

Cementum protein 1. CEMP1, recently identified as an extracellular matrix protein, is considered as a marker of cementum [14]. It plays a role in the differentiation of cementoblasts, and in the mineralization process by promoting the deposition and growth of hydroxyapatite crystals. It may have biological importance for tissue regeneration in periodontal disease. To our knowledge, no CEMP1 polymorphisms, including the variants R15Q, K55E and R80H detected in Denisova (Table 2), have been described associated with a particular phenotype. It can be noticed that the Denisova individual carries 3 substitutions (R15Q, K55E, R80H) that could be considered as a haplotype. K55E and R80H are located in the domain of the protein showing similarity with collagen alpha1 (I) [14] and all 3 mutations cause a change in the charge of the residue that could impact functional interactions. Noteworthy, K55E

corresponds to the ancestral variant, whereas the major substitution K55 in living humans is derived and nearly fixed out of Africa (Fig 3B), and likely represents a polymorphism specific to the modern lineage.

Matrix Gla protein. MGP is an essential extracellular matrix protein that exerts an inhibitory control on tissue calcification [14, 64]. Functional deficit or reduced expression of the protein lead to tissue over-mineralization. As MGP is expressed in kidney, lung and heart, in addition to bone and teeth, protein disfunction is associated with severe cardiovascular diseases. The nowadays common MGP variant T127A (rs4236) found widespread in the sampled archaic hominins (Table 2) was not associated with subgingival dental calculus [65] but was shown to decrease by 1.4–1.8 fold the risk of kidney stones in Han Chinese or Japanese patients [64, 66]. The T127A substitution suppresses a potential phosphorylation site at the C-terminal end of the protein that could be important for protein activity or for its interaction with other proteins or substrates. The other archaic and derived polymorphism A8V (rs142330429), nowadays restricted at low frequency in Asia and Finland, is not specifically found related to any relevant phenotype, although MGP is found among the 22 proteins inferred to be introgressed from archaic hominins in living humans, and which were sorted in relation to hypoplasia and skeleton abnormalities by phenotype-ontology enrichment tests [44].

Discussion

The analysis of tooth proteins in archaic hominins reveals 16 non-synonymous polymorphisms in 6 non-collagenous extracellular matrix proteins. Most of them correspond to derived variants for which the ancestral amino acid (common with apes) is in contrast present at high frequency in living people. This is in agreement with the global exome analysis of Castellano et al. [44] showing a higher proportion of derived amino acid changes in archaic compared to living humans and a greater susceptibility to altered protein structure and function. Considering all the Neanderthal missense derived variants, their frequencies are below 10% in living humans [67], which compares with the frequencies observed here for the archaic tooth variants (Table 2). A number of 274 missense polymorphisms for which the ancestral allele is in contrast fixed in modern humans is reported by Prufer et al. [46], and a minimum of 82 derived substitutions in 78 proteins can be estimated from Castellano et al. who sorted the variants by phenotype-ontology enrichment [44]. They found in particular, 10 Neanderthal-specific variants in 6 proteins involved in hyperlordosis, and 26 archaic-specific variants in 22 proteins associated in part with hypoplasia and skeleton abnormalities. Noteworthy, among them the MGP variant A8V certainly deserves to be considered in future genome wide association studies including dentition, and in the context of particular regional distribution (see below).

Most of the amino acid changes are found in the dentin and enamel proteins that are encoded by genes located in two close clusters on the chromosome 4 (Fig 4) and supposed to derive from an ancestral gene encoding a secretory calcium-binding protein [68]. These clusters fall apart from known haplotypes on chromosome 4 and do not correspond to candidate Neanderthal or Denisovan gene flow regions in non-African living humans [69, 70]. However, a signal of archaic introgression in the locus 4qMB179 on chromosome 4 has been evidenced in Central African populations, and is supposed to result from admixture with an archaic *Homo* some 35 kya [71]. In this study, the authors analyzed a non-coding intergenic region (>31 kb) within the locus. It would be therefore relevant to test if the region of the chromosome 4 that harbors the dental protein polymorphisms of Neanderthal and Denisova hominins may also reflect the signature of an archaic introgression in Africa. This is even more pertinent

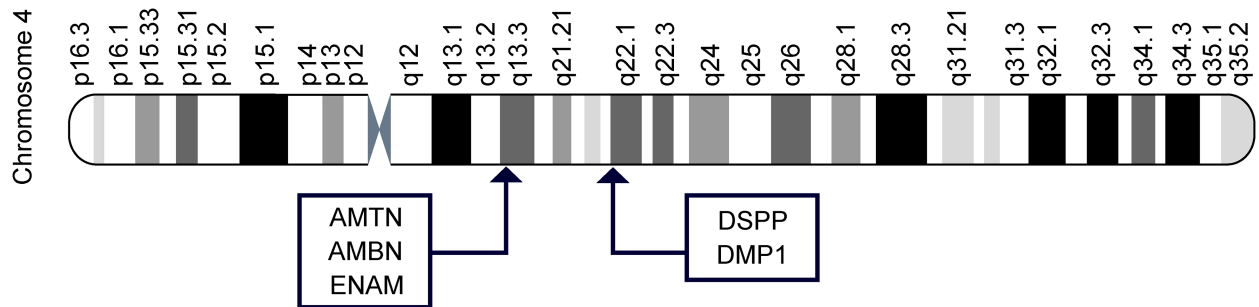


Fig 4. Localization of tooth protein genes in two clusters on chromosome 4.

<https://doi.org/10.1371/journal.pone.0183802.g004>

since there is evidence, as mentioned above, of positive selection of the ENAM gene [15–17] and since some other non-archaic missense variants in AMBN, ENAM and DMP1 are specifically expressed in Africa (S1 Fig).

Among the 16 archaic polymorphisms identified, only 2 are ancestral K55E (rs13331643) in CEMP1 and T127A (rs4236) in MGP. The CEMP1 variant K55E (rs13331643) is nowadays mostly restricted to Africa with a frequency around 25%, while in the rest of the world the derived modern-specific substitution K55 is the major allele. This suggests that the ancestral variant may have not been favorable to AMH outside Africa or, conversely, that the derived allele was specific to AMH and evolved out of Africa due to a founder effect. The ancestral archaic variant could also explain in part the tendency of people with African origins to have larger crowns and root volume and surface [2, 3, 52, 72]. Indeed, this variant is also found in Neanderthals teeth, characterized by high crown dimensions and frequent taurodontism [73]. The presence of this variant in Denisova is also compatible with such traits. Even if only two molars are reported for this taxon, they have extremely large crowns and Denisova 4 (the only one with fully preserved roots) displays massive diverging roots [33]. The MGP ancestral variant T127A (rs4236) is common (39% frequency) but less frequent in East Asia (around 10% frequency) suggesting that this ancestral allele was not adapted to the regional context, or was more rapidly diluted by admixture with local non-carrier population.

Other than the common worldwide derived ENAM polymorphism T648I (rs7671281) involved in a thinner enamel that is nearly fixed out of Africa and could be considered as *Homo*-specific, all derived polymorphisms are minor or rare alleles found restricted to either Africa or East/South Asia. Their frequencies fall in the range (< 0.1%) of the majority of the other non-archaic tooth protein missense variants (S1 Table). Derived variants with the highest frequencies do not exhibit specific regional distribution (S1 Fig) but, as for low and very low archaic variants, some polymorphisms in AMBN, ENAM, DMP1 and MGP, are distributed in either Africa or Asia (S1 Fig). Compared to Neanderthal missense variants located in proteins inferred to be introgressed [44, 67, 74], the frequencies of archaic tooth polymorphisms are lower (S2 Fig), indicating that the archaic derived variants in tooth proteins were probably not under positive selection, but may reflect a signature of population bottleneck in some regions and could trace persistence of archaic origin in specific populations.

The finding in living Africans of rare derived polymorphisms that were present in Neanderthals (Fig 2: ENAM G389S, 5%) or Denisova (Fig 3: CEMP1 R80H, 0.5%) is difficult to reconcile with the fact that interbreeding of AMH with Neanderthals occurred out of Africa [75, 76]. However, as mentioned above, this could be explained by the inside Africa admixture between the ancestors of AMH and archaic hominins, or by the possible back migration of Neanderthal gene flux [71, 75]. Further genome analysis of the Neanderthal and Denisova polymorphisms

in AMBN, ENAM, and CEMP1 certainly deserves attention as they could trace back an African ancestor common to Neanderthals, Denisova and early anatomically modern humans.

The other archaic derived variants (M273V in AMBN, R171H and A200V in AMTN, R15Q in CEMP1, N483T and R173Q in DMP1, A8V in MGP) are mainly found expressed at frequency between 0.1% to 4% in East and South Asia today, in accordance with highest proportion of Neanderthal and Denisova genomes remaining in their genomes and in the genome of South Pacific people [47, 70, 77]. Those variants could be responsible for tooth characteristics particular to Asian living people. Although there seems to be little variation in average and relative enamel thickness in present-day humans [78], Asians do show more complex dental traits than other populations. The main characteristic is a high prevalence of shoveling and double shoveling of the upper incisors that is less frequent in Europeans and Africans, additional cusp 6 and extra distal roots [52, 79]. According to [1], the highest frequencies are found in North Asia and the Americas. The lowest are in Western Eurasia and Sahul-Pacific. For the double shoveling, the highest frequencies are present among Native Americans (55–70%) and North Asia (62.5%) [1, 3]. The lowest are found in Southern Africa (0%) and sub-Saharan Africans (1.3%). Shovel-shaped incisors and crown size have been associated with an Asian-specific non-synonymous polymorphism in the Ectodysplasin receptor gene (EDAR) involved in hair thickness but absent in European, African and Denisovan populations [43, 79]. Interestingly, the shoveling phenotype is observed in *Homo erectus* and Neanderthal teeth [28, 79]. The EDAR variant does not explain all the Asian dental traits, and other genetic factors are likely involved, in particular regarding the development of the hypoconulid [43]. In support, the post-transcriptional repressor miR-1304 ancestral allele present in Neanderthals and today restricted to Asia (6% frequency in South Asia), which is responsible for the lower expression of the dental proteins enamelin and amelotin, may also shape the dentition or contribute to a higher risk of amelogenesis imperfecta in Asian populations [26]. In this regard, the polymorphisms common to modern East/South Asians and to Neanderthal or Denisova individuals, identified here in AMTN, DMP1 and MGP, could also have an influence on tooth structure or dental health in these populations. Investigating the functional consequences of these variants in relation to the prevalence of dental and other diseases in the Asian population certainly deserves interest, in particular for A8V in the protein MGP that is found associated in phenotype-ontology test with skeleton abnormalities, cardiovascular and metabolism diseases, hair and morphological traits [44].

In conclusion, this study presents a list of archaic tooth proteins variants with potential matter of investigation for those who are interested in the relationships between protein polymorphisms and dental morphostructure, taxonomic determination of archaic lineages, or characterization of particular traits or disease susceptibility in diverse present-day populations [72]. However, further biochemical and functional exploration or genome wide association studies of the polymorphisms are now required. The present study also points to the specific regional distribution of rare dental polymorphisms with archaic origin around the world, bringing some support to the geographic morphological patterns described earlier [1, 3]. Deciphering the functional impact of these polymorphisms on teeth could be relevant for dental practitioners to adapt their treatment in view of personalized medicine. It also illustrates that our knowledge of the functional consequence of the human polymorphism diversity is essentially based on population with European ancestry, and that informative data from other ethnic groups should also be included [80].

Supporting information

S1 Table. Tooth protein missense variants with a frequency > 0.001 (1000 Genome or ExAC, <http://www.ensembl.org>). Archaic variants are indicated in bold characters. AFR:

Africa AMR: America EUR: Europe EAS: East Asia SAS: South Asia.
(PDF)

S1 Fig. Geographical distribution of non-synonymous variants (frequency > 0.001) in tooth proteins. For clarity in the ENAM diagram, the major variant T648I was not considered.

* Archaic variants.

(PDF)

S2 Fig. Frequency of archaic missense variants in tooth proteins compared to proteins inferred to be introgressed. ¹CUL7, HSPG2, NEB, FRAS1, FREM2 are found in Neanderthal or archaic (Neanderthal + Denisova) introgressed regions (Castellano et al. 2014). ²ZNF365 and SLC16A1 are found associated with Crohn's disease in Ashkenazi Jews and Type II diabetes in Latinos, respectively (Sankararaman et al., 2014). ³SIPA1L2 is present in haplotype with signature of adaptive introgression in Asia (Vernot and Akey, 2014). # Ancestral variants.

(PDF)

Acknowledgments

We are very grateful to M Gibert, P Balaesque, N Brucato and E Crubézy (AMIS, Toulouse) and J-S Saulnier-Blache (I2MC, Toulouse) for helpful discussion and comments during the preparation and the revision of the manuscript. We also thank X Mata (AMIS, Toulouse) for advice on bioinformatics. The work was supported by CNRS (PEPS blanc 2016) and IFRO (2016). The present version greatly benefited from the comments provided by the Editors and three anonymous reviewers.

Author Contributions

Conceptualization: Clément Zanolli, Catherine Mollereau.

Formal analysis: Mathilde Hourset, Catherine Mollereau.

Funding acquisition: Rémi Esclassan.

Investigation: Catherine Mollereau.

Methodology: Mathilde Hourset, Catherine Mollereau.

Project administration: Catherine Mollereau.

Supervision: Catherine Mollereau.

Visualization: Rémi Esclassan.

Writing – original draft: Clément Zanolli, Catherine Mollereau.

Writing – review & editing: Clément Zanolli, Catherine Mollereau.

References

1. Bailey SE, Hublin JJ. What does it mean to be dentally "modern"? *Anthropological Perspectives on Tooth Morphology: Genetics, Evolution, Variation* p222-249. Scott G, Irish J, editors: Cambridge University Press; 2013.
2. Harris EF, Hicks JD, Barcroft BD. Tissue contributions to sex and race: differences in tooth crown size of deciduous molars. *Am J Phys Anthropol.* 2001; 115(3):223–37. <https://doi.org/10.1002/ajpa.1077> PMID: 11424074.
3. Scott RG, Turner CG. Dental morphology and its variation in recent human populations. *The anthropology of modern human teeth* p1-371: Cambridge University Press; 1997.

4. Bei M. Molecular genetics of tooth development. *Curr Opin Genet Dev.* 2009; 19(5):504–10. <https://doi.org/10.1016/j.gde.2009.09.002> PMID: 19875280;
5. Harjunmaa E, Seidel K, Hakkinen T, Renvoise E, Corfe IJ, Kallonen A, et al. Replaying evolutionary transitions from the dental fossil record. *Nature.* 2014; 512(7512):44–8. <https://doi.org/10.1038/nature13613> PMID: 25079326;
6. Jernvall J, Thesleff I. Reiterative signaling and patterning during mammalian tooth morphogenesis. *Mech Dev.* 2000; 92(1):19–29. PMID: 10704885.
7. Salazar-Ciudad I, Jernvall J. A gene network model accounting for development and evolution of mammalian teeth. *Proc Natl Acad Sci U S A.* 2002; 99(12):8116–20. <https://doi.org/10.1073/pnas.132069499> PMID: 12048258;
8. Suryadeva S, Khan MB. Role of homeobox genes in tooth morphogenesis: a review. *J Clin Diagn Res.* 2015; 9(2):ZE09–12. <https://doi.org/10.7860/JCDR/2015/11067.5606> PMID: 25859538;
9. Kato A, Tang N, Borries C, Papakyrikos AM, Hinde K, Miller E, et al. Intra- and interspecific variation in macaque molar enamel thickness. *Am J Phys Anthropol.* 2014; 155(3):447–59. <https://doi.org/10.1002/ajpa.22593> PMID: 25146639.
10. Pampush JD, Duque AC, Burrows BR, Daegling DJ, Kenney WF, McGraw WS. Homoplasy and thick enamel in primates. *Journal of human evolution.* 2013; 64(3):216–24. <https://doi.org/10.1016/j.jhevol.2013.01.009> PMID: 23415414.
11. Jagr M, Eckhardt A, Pataridis S, Broukal Z, Duskova J, Miksik I. Proteomics of human teeth and saliva. *Physiol Res.* 2014; 63 Suppl 1:S141–54. PMID: 24564654.
12. Seow WK. Developmental defects of enamel and dentine: challenges for basic science research and clinical management. *Aust Dent J.* 2014; 59 Suppl 1:143–54. <https://doi.org/10.1111/adj.12104> PMID: 24164394.
13. Castiblanco GA, Rutishauser D, Ilag LL, Martignon S, Castellanos JE, Mejia W. Identification of proteins from human permanent erupted enamel. *Eur J Oral Sci.* 2015; 123(6):390–5. <https://doi.org/10.1111/eos.12214> PMID: 26432388.
14. Arzate H, Zeichner-David M, Mercado-Celis G. Cementum proteins: role in cementogenesis, biomineralization, periodontium formation and regeneration. *Periodontol 2000.* 2015; 67(1):211–33. <https://doi.org/10.1111/prd.12062> PMID: 25494602.
15. Daubert DM, Kelley JL, Udod YG, Habor C, Kleist CG, Furman IK, et al. Human enamel thickness and ENAM polymorphism. *Int J Oral Sci.* 2016; 8(2):93–7. <https://doi.org/10.1038/ijos.2016.1> PMID: 27357321;
16. Horvath JE, Ramachandran GL, Fedrigo O, Nielsen WJ, Babbitt CC, St Clair EM, et al. Genetic comparisons yield insight into the evolution of enamel thickness during human evolution. *Journal of human evolution.* 2014; 73:75–87. <https://doi.org/10.1016/j.jhevol.2014.01.005> PMID: 24810709.
17. Kelley JL, Swanson WJ. Dietary change and adaptive evolution of enamelin in humans and among primates. *Genetics.* 2008; 178(3):1595–603. <https://doi.org/10.1534/genetics.107.077123> PMID: 18245370;
18. Macchiarelli R, Mazurier A, Illerhaus B, Zanolli C. *Ouranopithecus macedoniensis* (Mammalia, Primates, Hominoidea): virtual reconstruction and 3D analysis of a juvenile mandibular dentition (RPI-82 and RPI-83). *Geodiversitas.* 2009; 31(4):851–63. <https://doi.org/10.5252/g2009n4a851>
19. Olejniczak AJ, Smith TM, Skinner MM, Grine FE, Feeney RN, Thackeray JF, et al. Three-dimensional molar enamel distribution and thickness in *Australopithecus* and *Paranthropus*. *Biol Lett.* 2008; 4(4):406–10. <https://doi.org/10.1098/rsbl.2008.0223> PMID: 18522924;
20. Pan L, Dumoncel J, de Beer F, Hoffman J, Thackeray JF, Duployer B, et al. Further morphological evidence on South African earliest *Homo* lower postcanine dentition: Enamel thickness and enamel dentine junction. *Journal of human evolution.* 2016; 96:82–96. Epub 2016/06/28. <https://doi.org/10.1016/j.jhevol.2016.05.003> PMID: 27343773.
21. Skinner MM, Wood BA, Boesch C, Olejniczak AJ, Rosas A, Smith TM, et al. Dental trait expression at the enamel-dentine junction of lower molars in extant and fossil hominoids. *Journal of human evolution.* 2008; 54(2):173–86. <https://doi.org/10.1016/j.jhevol.2007.09.012> PMID: 18048081.
22. Smith TM, Olejniczak AJ, Zermeno JP, Tafforeau P, Skinner MM, Hoffmann A, et al. Variation in enamel thickness within the genus *Homo*. *Journal of human evolution.* 2012; 62(3):395–411. <https://doi.org/10.1016/j.jhevol.2011.12.004> PMID: 22361504.
23. Zanolli C. Molar crown inner structural organization in Javanese *Homo erectus*. *Am J Phys Anthropol.* 2015; 156(1):148–57. Epub 2014/09/12. <https://doi.org/10.1002/ajpa.22611> PMID: 25209431.
24. Olejniczak AJ, Smith TM, Feeney RN, Macchiarelli R, Mazurier A, Bondioli L, et al. Dental tissue proportions and enamel thickness in Neandertal and modern human molars. *Journal of human evolution.* 2008; 55(1):12–23. <https://doi.org/10.1016/j.jhevol.2007.11.004> PMID: 18321561.

25. Zanolli C, Bondioli L, Coppa A, Dean CM, Bayle P, Candilio F, et al. The late Early Pleistocene human dental remains from Uadi Aalad and Mulhuli-Amo (Buia), Eritrean Danakil: Macromorphology and microstructure. *Journal of human evolution*. 2014; 74:96–113. <https://doi.org/10.1016/j.jhevol.2014.04.005> PMID: 24852385
26. Lopez-Valenzuela M, Ramirez O, Rosas A, Garcia-Vargas S, de la Rasilla M, Lalueza-Fox C, et al. An ancestral miR-1304 allele present in Neanderthals regulates genes involved in enamel formation and could explain dental differences with modern humans. *Mol Biol Evol*. 2012; 29(7):1797–806. <https://doi.org/10.1093/molbev/mss023> PMID: 22319171.
27. Macchiarelli R, Bondioli L, Debenath A, Mazurier A, Tournepiche JF, Birch W, et al. How Neanderthal molar teeth grew. *Nature*. 2006; 444(7120):748–51. <https://doi.org/10.1038/nature05314> PMID: 17122777.
28. Martinon-Torres M, Bermudez de Castro JM, Gomez-Robles A, Arsuaga JL, Carbonell E, Lordkipanidze D, et al. Dental evidence on the hominin dispersals during the Pleistocene. *Proc Natl Acad Sci U S A*. 2007; 104(33):13279–82. <https://doi.org/10.1073/pnas.0706152104> PMID: 17684093;
29. Martin RM, Hublin JJ, Gunz P, Skinner MM. The morphology of the enamel-dentine junction in Neanderthal molars: Gross morphology, non-metric traits, and temporal trends. *Journal of human evolution*. 2017; 103:20–44. <https://doi.org/10.1016/j.jhevol.2016.12.004> PMID: 28166906.
30. Smith TM, Tafforeau P, Reid DJ, Pouech J, Lazzari V, Zermeno JP, et al. Dental evidence for ontogenetic differences between modern humans and Neanderthals. *Proc Natl Acad Sci U S A*. 2010; 107(49):20923–8. <https://doi.org/10.1073/pnas.1010906107> PMID: 21078988;
31. de Pinillos MM, Martinon-Torres M, Skinner MM, Arsuaga JL, Gracia-Tellez A, Martinez I, et al. Trigonid crests expression in Atapuerca-Sima de los Huesos lower molars: Internal and external morphological expression and evolutionary inferences. *Cr Palevol*. 2014; 13(3):205–21. <https://doi.org/10.1016/j.crpv.2013.10.008>
32. Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*. 2010; 468(7327):1053–60. Epub 2010/12/24. <https://doi.org/10.1038/nature09710> PMID: 21179161;
33. Sawyer S, Renaud G, Viola B, Hublin JJ, Gansauge MT, Shunkov MV, et al. Nuclear and mitochondrial DNA sequences from two Denisovan individuals. *Proc Natl Acad Sci U S A*. 2015; 112(51):15696–700. Epub 2015/12/03. <https://doi.org/10.1073/pnas.1519905112> PMID: 26630009;
34. Guy F, Lazzari V, Gilissen E, Thiery G. To What Extent is Primate Second Molar Enamel Occlusal Morphology Shaped by the Enamel-Dentine Junction? *PLoS One*. 2015; 10(9):e0138802. <https://doi.org/10.1371/journal.pone.0138802> PMID: 26406597;
35. Chaussain C, Bouazza N, Gasse B, Laffont AG, Opsahl Vital S, Davit-Beal T, et al. Dental caries and enamelin haplotype. *J Dent Res*. 2014; 93(4):360–5. <https://doi.org/10.1177/0022034514522060> PMID: 24487377.
36. Gadhia K, McDonald S, Arkutu N, Malik K. Amelogenesis imperfecta: an introduction. *Br Dent J*. 2012; 212(8):377–9. <https://doi.org/10.1038/sj.bdj.2012.314> PMID: 22538897.
37. Hart TC, Hart PS, Gorry MC, Michalec MD, Ryu OH, Uygun C, et al. Novel ENAM mutation responsible for autosomal recessive amelogenesis imperfecta and localised enamel defects. *J Med Genet*. 2003; 40(12):900–6. <https://doi.org/10.1136/jmg.40.12.900> PMID: 14684688;
38. McKnight DA, Suzanne Hart P, Hart TC, Hartsfield JK, Wilson A, Wright JT, et al. A comprehensive analysis of normal variation and disease-causing mutations in the human DSPP gene. *Hum Mutat*. 2008; 29(12):1392–404. <https://doi.org/10.1002/humu.20783> PMID: 18521831.
39. Opal S, Garg S, Jain J, Walia I. Genetic factors affecting dental caries risk. *Aust Dent J*. 2015; 60(1):2–11. <https://doi.org/10.1111/adj.12262> PMID: 25721273.
40. Poulter JA, Murillo G, Brookes SJ, Smith CE, Parry DA, Silva S, et al. Deletion of ameloblastin exon 6 is associated with amelogenesis imperfecta. *Hum Mol Genet*. 2014; 23(20):5317–24. <https://doi.org/10.1093/hmg/ddu247> PMID: 24858907;
41. Smith CE, Murillo G, Brookes SJ, Poulter JA, Silva S, Kirkham J, et al. Deletion of amelotin exons 3–6 is associated with amelogenesis imperfecta. *Hum Mol Genet*. 2016; 25(16):3578–87. <https://doi.org/10.1093/hmg/ddw203> PMID: 27412008;
42. Wright JT, Hart TC, Hart PS, Simmons D, Suggs C, Daley B, et al. Human and mouse enamel phenotypes resulting from mutation or altered expression of AMEL, ENAM, MMP20 and KLK4. *Cells Tissues Organs*. 2009; 189(1–4):224–9. <https://doi.org/10.1159/000151378> PMID: 18714142;
43. Park JH, Yamaguchi T, Watanabe C, Kawaguchi A, Haneji K, Takeda M, et al. Effects of an Asian-specific nonsynonymous EDAR variant on multiple dental traits. *J Hum Genet*. 2012; 57(8):508–14. <https://doi.org/10.1038/jhg.2012.60> PMID: 22648185.

44. Castellano S, Parra G, Sanchez-Quinto FA, Racimo F, Kuhlwil M, Kircher M, et al. Patterns of coding variation in the complete exomes of three Neandertals. *Proc Natl Acad Sci U S A*. 2014; 111(18):6666–71. Epub 2014/04/23. <https://doi.org/10.1073/pnas.1405138111> PMID: 24753607;
45. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science*. 2012; 338(6104):222–6. Epub 2012/09/01. <https://doi.org/10.1126/science.1224344> PMID: 22936568;
46. Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*. 2014; 505(7481):43–9. Epub 2013/12/20. <https://doi.org/10.1038/nature12886> PMID: 24352235;
47. Nielsen R, Akey JM, Jakobsson M, Pritchard JK, Tishkoff S, Willerslev E. Tracing the peopling of the world through genomics. *Nature*. 2017; 541(7637):302–10. <https://doi.org/10.1038/nature21347> PMID: 28102248.
48. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome research*. 2002; 12(6):996–1006. Epub 2002/06/05. Article published online before print in May 2002. <https://doi.org/10.1101/gr.229102> PMID: 12045153;
49. Herrero J, Muffato M, Beal K, Fitzgerald S, Gordon L, Pignatelli M, et al. Ensembl comparative genomics resources. *Database (Oxford)*. 2016; 2016. Epub 2016/02/21. <https://doi.org/10.1093/database/bav096> PMID: 26896847;
50. Marcus JH, Novembre J. Visualizing the geography of genetic variants. *Bioinformatics*. 2016. <https://doi.org/10.1093/bioinformatics/btw643> PMID: 27742697.
51. Meyer M, Arsuaga JL, de Filippo C, Nagel S, Aximu-Petri A, Nickel B, et al. Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature*. 2016; 531(7595):504–7. <https://doi.org/10.1038/nature17405> PMID: 26976447.
52. Hanihara T. Morphological variation of major human populations based on nonmetric dental traits. *Am J Phys Anthropol*. 2008; 136(2):169–82. <https://doi.org/10.1002/ajpa.20792> PMID: 18257017.
53. Skinner MM, Alemseged Z, Gaunitz C, Hublin JJ. Enamel thickness trends in Plio-Pleistocene hominin mandibular molars. *Journal of human evolution*. 2015; 85:35–45. <https://doi.org/10.1016/j.jhevol.2015.03.012> PMID: 26024565.
54. Arnaud J, Benazzi S, Romandini M, Livraghi A, Panetta D, Salvadori PA, et al. A Neanderthal deciduous human molar with incipient carious infection from the Middle Palaeolithic De Nadale cave, Italy. *Am J Phys Anthropol*. 2017; 162(2):370–6. <https://doi.org/10.1002/ajpa.23111> PMID: 27704536.
55. Zanolli C, Mazurier A. Endostructural characterization of the H. heidelbergensis dental remains from the early Middle Pleistocene site of Tighenif, Algeria. *Cr Palevol*. 2013; 12(5):293–304. <https://doi.org/10.1016/j.crpv.2013.06.004>
56. Zanolli C, Dean MC, Assefa Y, Bayle P, Braga J, Condemi S, et al. Structural organization and tooth development in a Homo aff. erectus juvenile mandible from the Early Pleistocene site of Garba IV at Melka Kunture, Ethiopian highlands. *Am J Phys Anthropol*. 2017; 162(3):533–49. <https://doi.org/10.1002/ajpa.23135> PMID: 27883188.
57. Zilberman U, Smith P, Piperno M, Condemi S. Evidence of amelogenesis imperfecta in an early African Homo erectus. *Journal of human evolution*. 2004; 46(6):647–53. <https://doi.org/10.1016/j.jhevol.2004.02.005> PMID: 15183668.
58. Al-Hashimi N, Sire JY, Delgado S. Evolutionary analysis of mammalian enamelin, the largest enamel protein, supports a crucial role for the 32-kDa peptide and reveals selective adaptation in rodents and primates. *J Mol Evol*. 2009; 69(6):635–56. <https://doi.org/10.1007/s00239-009-9302-x> PMID: 20012271.
59. Chan HC, Mai L, Oikonomopoulou A, Chan HL, Richardson AS, Wang SK, et al. Altered enamelin phosphorylation site causes amelogenesis imperfecta. *J Dent Res*. 2010; 89(7):695–9. <https://doi.org/10.1177/0022034510365662> PMID: 20439930;
60. Silvent J, Sire JY, Delgado S. The dentin matrix acidic phosphoprotein 1 (DMP1) in the light of mammalian evolution. *J Mol Evol*. 2013; 76(1–2):59–70. <https://doi.org/10.1007/s00239-013-9539-2> PMID: 23361408.
61. Smith CE, Wazen R, Hu Y, Zalzal SF, Nanci A, Simmer JP, et al. Consequences for enamel development and mineralization resulting from loss of function of ameloblastin or enamelin. *Eur J Oral Sci*. 2009; 117(5):485–97. <https://doi.org/10.1111/j.1600-0722.2009.00666.x> PMID: 19758243;
62. Delsuc F, Gasse B, Sire JY. Evolutionary analysis of selective constraints identifies ameloblastin (AMBN) as a potential candidate for amelogenesis imperfecta. *BMC Evol Biol*. 2015; 15:148. <https://doi.org/10.1186/s12862-015-0431-0> PMID: 26223266;

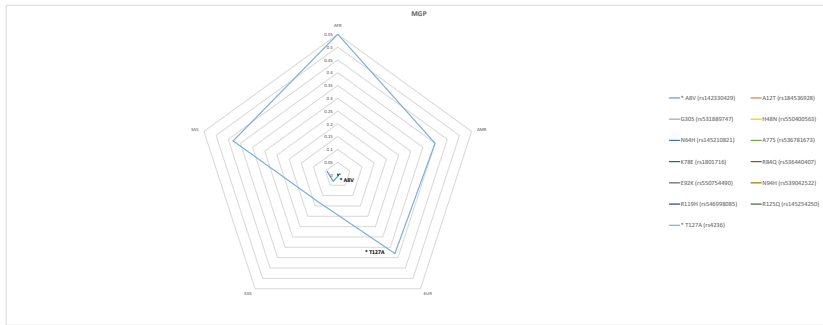
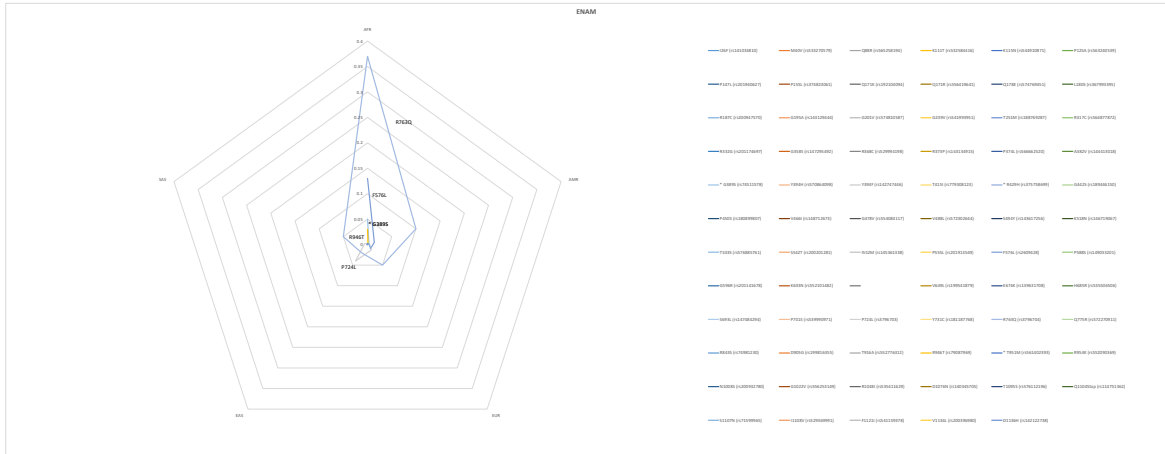
63. Gasse B, Silvent J, Sire JY. Evolutionary Analysis Suggests That AMTN is Enamel-specific and a Candidate for AI. *Journal of Dental Research*. 2012; 91(11):1085–9. <https://doi.org/10.1177/0022034512460551> PMID: 22968158
64. Gao B, Yasui T, Itoh Y, Tozawa K, Hayashi Y, Kohri K. A polymorphism of matrix Gla protein gene is associated with kidney stones. *J Urol*. 2007; 177(6):2361–5. <https://doi.org/10.1016/j.juro.2007.01.118> PMID: 17509359.
65. Dogan GE, Demir T, Aksoy H, Saglam E, Laloglu E, Yildirim A. Matrix-Gla Protein rs4236 [A/G] gene polymorphism and serum and GCF levels of MGP in patients with subgingival dental calculus. *Arch Oral Biol*. 2016; 70:125–9. <https://doi.org/10.1016/j.archoralbio.2016.06.014> PMID: 27348051.
66. Lu X, Gao B, Liu Z, Tian X, Mao X, Emmanuel N, et al. A polymorphism of matrix Gla protein gene is associated with kidney stone in the Chinese Han population. *Gene*. 2012; 511(2):127–30. <https://doi.org/10.1016/j.gene.2012.09.112> PMID: 23046575.
67. Sankararaman S, Mallick S, Dannemann M, Prufer K, Kelso J, Paabo S, et al. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*. 2014; 507(7492):354–7. Epub 2014/01/31. <https://doi.org/10.1038/nature12961> PMID: 24476815;
68. Yoshizaki K, Yamada Y. Gene evolution and functions of extracellular matrix proteins in teeth. *Orthod Waves*. 2013; 72(1):1–10. <https://doi.org/10.1016/j.odw.2013.01.040> PMID: 23539364;
69. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, et al. A draft sequence of the Neanderthal genome. *Science*. 2010; 328(5979):710–22. Epub 2010/05/08. <https://doi.org/10.1126/science.1188021> PMID: 20448178.
70. Vernot B, Tucci S, Kelso J, Schraiber JG, Wolf AB, Gittelman RM, et al. Excavating Neanderthal and Denisovan DNA from the genomes of Melanesian individuals. *Science*. 2016; 352(6282):235–9. <https://doi.org/10.1126/science.aad9416> PMID: 26989198.
71. Hammer MF, Woerner AE, Mendez FL, Watkins JC, Wall JD. Genetic evidence for archaic admixture in Africa. *Proc Natl Acad Sci U S A*. 2011; 108(37):15123–8. <https://doi.org/10.1073/pnas.1109300108> PMID: 21896735;
72. Alliot-Licht B, Lusson C, Hyon I, Dajeau-Trudaud S, Le Caignec C, Lopez-Cazaux S. [Extra-oral signs to look for in patients exhibiting oral warning signs of genetic diseases]. *C R Biol*. 2015; 338(1):48–57. <https://doi.org/10.1016/j.crv.2014.10.001> PMID: 25528675.
73. Bailey SE. A closer look at Neanderthal postcanine dental morphology: The mandibular dentition. *The Anatomical Record*. 2002; 269:148–56. <https://doi.org/10.1002/ar.10116> PMID: 12124901
74. Vernot B, Akey JM. Resurrecting surviving Neanderthal lineages from modern human genomes. *Science*. 2014; 343(6174):1017–21. <https://doi.org/10.1126/science.1245938> PMID: 24476670.
75. Hsieh P, Woerner AE, Wall JD, Lachance J, Tishkoff SA, Gutenkunst RN, et al. Model-based analyses of whole-genome data reveal a complex evolutionary history involving archaic introgression in Central African Pygmies. *Genome research*. 2016; 26(3):291–300. <https://doi.org/10.1101/gr.196634.115> PMID: 26888264;
76. Paabo S. The human condition—a molecular approach. *Cell*. 2014; 157(1):216–26. Epub 2014/04/01. <https://doi.org/10.1016/j.cell.2013.12.036> PMID: 24679537.
77. Sankararaman S, Mallick S, Patterson N, Reich D. The Combined Landscape of Denisovan and Neanderthal Ancestry in Present-Day Humans. *Curr Biol*. 2016; 26(9):1241–7. <https://doi.org/10.1016/j.cub.2016.03.037> PMID: 27032491;
78. Feeney RNM, Zermeno JP, Reid DJ, Nakashima S, Sano H, Bahar A, et al. Enamel thickness in Asian human canines and premolars. *Anthropol Sci*. 2010; 118(3):191–8. <https://doi.org/10.1537/ase.091006>
79. Kimura R, Yamaguchi T, Takeda M, Kondo O, Toma T, Haneji K, et al. A common variation in EDAR is a genetic determinant of shovel-shaped incisors. *Am J Hum Genet*. 2009; 85(4):528–35. <https://doi.org/10.1016/j.ajhg.2009.09.006> PMID: 19804850;
80. Lu YF, Goldstein DB, Angrist M, Cavalleri G. Personalized medicine and human genetic diversity. *Cold Spring Harb Perspect Med*. 2014; 4(9):a008581. <https://doi.org/10.1101/cshperspect.a008581> PMID: 25059740;

AMBN (over 209 missenses)	Allele frequency of the variant					
	All	AFR	AMR	EUR	EAS	SAS
S41N (rs546778141) derived	0.16	0.07	0.19	0.24	0.3	0.01
M76V (rs200620822) derived					0.001	
M76I (rs145190129) derived	0.002	0.01				
G78V (rs145763811) derived	0.005	0.016	0.006			
G78S (rs143795139) derived		0.0001				
S102Y (rs145745749) derived	0.006	0.02	0.006			
P125H (rs180819235) derived			0.001			
E142K (rs550317645) derived						0.001
V181A (rs189846864) derived	0.001		0.004			
D182Y (rs543637559) derived				0.001		
R222H (rs142755020) derived					0.001	
P230S (rs369135444) derived	0.001					0.003
V244M (rs201976575) derived	0.01					0.05
A255V (rs7439186) derived	0.12	0.15	0.073	0.065	0.15	0.14
G268R (rs146167261) derived	0.002			0.002		0.006
M273V (rs564905233) derived	0.001					0.003
G276R (rs201299436) derived					0.001	
M285I (rs115723025) derived	0.034	0.12	0.01			
G288S (rs568675718) derived	0.001					0.004
P293S (rs200796492) derived			0.001			
H294Q (rs113506649) derived	0.02	0.065	0.01			
G300S (rs139319140) derived	0.002	0.007			0.001	
E317A (rs150017698) derived	0.003	0.01	0.001			
M328T (rs538513467) derived			0.001			
P329L (rs140331879) derived	0.004		0.003		0.02	0.001
A351T (rs141401324) derived		0.001				
L354P (rs72654387) derived	0.02		0.024	0.06		0.01
G363S (rs543838539) derived					0.001	
V378I (rs76503327) derived	0.024	0.09	0.006			
L391S (rs141535122) derived				0.002		
T416A (rs527740107) derived		0.001				
T417M (rs146171297) derived						0.001
E436K (rs149937358) derived	0.001			0.002	0.003	0.001
A441T (rs144829376) derived	0.001	0.002		0.001		
AMTN (88 missenses)	Allele frequency of the variant					
All	AFR	AMR	EUR	EAS	SAS	
R15W (rs35286445) derived	0.11	0.007	0.11	0.16	0.19	0.14
P18S (rs35426053) derived	0.006	0.02	0.001			
P33L (rs199658784) derived						0.001
N45S (rs7660807) derived	0.1	0.006	0.06	0.1	0.2	0.17
V47F (rs192563032) derived	0.001	0.007				
S50P (rs34803339) derived	0.03	0.07	0.01		0.001	0.04
G78S (rs151041998) derived	0.001			0.004		0.001
G88E (rs146006787) derived	0.001		0.003	0.004		
L89M (rs560519245) derived						0.001
A109T (rs201205218) derived	0.001		0.001	0.001		0.003
H129Y (rs148416277) derived		0.002				
G166E (rs565052404) derived		0.001				
R171C (rs201349837) derived	0.001	0.002			0.001	
R171H (rs117133443) derived						
G177V (rs61755875) derived	0.02	0.06	0.004	0.01	0.005	
D181N (rs185210496) derived					0.001	
A200V (rs376682442) derived	0.004					0.02
CEMP1 (over 143 missenses)	Allele frequency of the variant					
All	AFR	AMR	EUR	EAS	SAS	
A11G (rs547680075) derived						0.001
R15Q (rs370735504) derived			0.001		0.001	
S19F (rs201463378) derived			0.001			
T27I (rs563820819) derived		0.001				

C28R (rs200449794) derived					0.001	
G50V (rs574778073) derived					0.001	
K55E (rs13331643) ancestral	0.07	0.26	0.02	0.002		0.001
R80H (rs79379654) derived	0.001	0.005				
R81H (rs200977329) derived			0.003			
A103S (rs535834483) derived	0.002					0.008
A107G (rs76722821) derived	0.008		0.02	0.03		0.002
R113K (rs202046694) derived			0.001			
S120P (rs200124887) derived		0.001				
P164L (rs531558523) derived						0.001
V165I (rs563584391) derived						0.001
P173Q (rs543906255) derived		0.001				
V182L (rs189095388) derived				0.002		
P214L (rs185302486) derived					0.001	
R216Q (rs141294706) derived		0.002				
T223M (rs558847727) derived						0.001
R225C (rs181209921) ancestral					0.001	
R225H (rs576590180) derived			0.001			
G234S (rs536394645) derived			0.001			
T238K (rs61746888) derived		0.001		0.001		
DMP1 (over 216 missenses)	Allele frequency of the variant					
	All	AFR	AMR	EUR	EAS	SAS
R21K (rs552859119) derived				0.001		
S27P (rs150105108) derived					0.001	
T42P (rs369550864) derived	0.001				0.004	
E49G (rs200704471) derived		0.001				
E60D (rs546062219) derived						0.001
S69C (rs10019009) derived	0.29	0.26	0.33	0.23	0.43	0.21
D80N (rs567911405) derived			0.001			
A88V (rs144773084) derived		0.001				
F91L (rs556512826) derived						0.001
T113I (rs145164698) derived		0.002				
D141N (rs144580319) derived	0.001	0.002				
D142Y (rs201761802) derived	0.002				0.008	
Q159K (rs79402270) derived	0.016		0.04	0.003	0.05	0.002
R173Q (rs747698893) derived					0.0001	0.001
T184I (rs149938637) derived				0.001		
G207D (rs200520896) derived				0.001		
S225N (rs373051924) derived			0.001			
M237V (rs201413886) derived					0.001	
S242P (rs147552663) derived		0.002				
S246R (rs190314299) derived			0.001			
P263A (rs571820459) derived				0.001		
R272H (rs145237146) derived	0.001		0.004	0.003		
L282I (rs141979823) derived	0.003		0.004		0.01	
M288T (rs565811736) derived		0.001				
E329D (rs554234835) derived				0.001		
V333I (rs147275271) derived	0.002	0.005				0.001
E362Q (rs558823933) derived				0.001		
E362A (rs576618731) derived				0.001		
D369N (rs540873212) derived					0.001	
S384N (rs148498977) derived		0.001				
S389L (rs572239643) derived		0.001				
E402K (rs185655763) derived				0.001		
P419S (rs140275311) derived				0.001		
P419R (rs145278609) derived		0.001				
S427T (rs532733390) derived					0.001	
K463R (rs34661425) derived	0.02	0.06	0.01	0.002		
T469M (rs199734764) derived					0.001	
D478A (rs148156611) derived	0.001		0.003	0.001		
N483T (rs574215585) derived	0.004	0.001				0.02
N498S (rs373096835) derived	0.001	0.002				

I501T (rs146762807) derived	0.001	0.002				0.001
ENAM (over 494 missenses)	Allele frequency of the variant					
	All	AFR	AMR	EUR	EAS	SAS
I26F (rs141034810) derived	0.001				0.005	
M40V (rs533270579) derived		0.001				
Q88R (rs565258194) derived			0.001			
K111T (rs532584416) derived						0.001
K115N (rs544910871) derived						0.001
P125A (rs563240539) derived						0.002
P147L (rs201940627) derived			0.001			
P155L (rs374823061) derived		0.001				
Q171K (rs192104094) derived					0.002	
Q171R (rs556419641) derived				0.001		
Q178E (rs574769451) derived		0.001				
L180S (rs367993395) derived		0.002				
R187C (rs200947570) derived			0.001			
G195A (rs143129444) derived		0.001				0.001
G201V (rs574810587) derived						0.001
G239V (rs541993951) derived					0.002	
T251M (rs188769287) derived					0.001	0.001
R317C (rs564877872) derived						0.001
R332G (rs201174697) derived		0.001				
G358S (rs147295492) derived	0.001				0.003	
R368C (rs529994198) derived					0.001	
R373P (rs143134915) derived						0.001
P374L (rs566662520) derived				0.001		
A382V (rs144413018) derived					0.001	
G389S (rs74511578) derived	0.013	0.05	0.003			
Y394H (rs570864098) derived					0.002	
Y394F (rs142747446) derived	0.001			0.002		0.002
T415I (rs779308123) derived				< 0.0001		
R429H (rs375758699) derived	0.0005					
G442S (rs189446150) derived			0.003			
P450S (rs180899807) derived					0.002	
V466I (rs148712673) derived	0.002	0.001		0.007		
G478V (rs554084117) derived						0.001
V488L (rs572302644) derived		0.001				
S494Y (rs143617256) derived				0.001		
K518N (rs146719067) derived				0.001		
T533S (rs576885761) derived					0.002	
S542T (rs200201281) derived				0.002		
I552M (rs145361338) derived						0.001
P555L (rs201913549) derived					0.001	
F576L (rs2609428) derived	0.04	0.13	0.014	0.01		
P588S (rs149033201) derived		0.002				
G596R (rs201141678) derived						0.001
K633N (rs552101482) derived			0.001			
T648I (rs7671281) derived	0.82	0.47	0.89	0.95	0.98	0.94
V649L (rs199541879) derived				0.001		
E674K (rs139631708) derived					0.001	
H685R (rs535506506) derived			0.001			
S693L (rs147484294) derived			0.001			
P701S (rs539990971) derived			0.001			
P724L (rs3796703) derived	0.013	0.002	0.004	0.013	0.04	0.006
Y731C (rs181187768) derived	0.001					
R763Q (rs3796704) derived	0.14	0.37	0.1	0.05	0.02	0.05
Q775R (rs572270911) derived						0.001
R843S (rs74981230) derived	0.002				0.01	
D905G (rs199816355) derived				0.001		
T916A (rs552776312) derived						0.002
R946T (rs79087969) derived	0.007	0.03	0.001			
T951M (rs561402393) derived					0.0001	0.0001

R954K (rs552090369) derived				0.001		
N1008S (rs200932780) derived				0.001		
G1022V (rs556253149) derived						0.001
R1048I (rs535411629) derived			0.001			
D1076N (rs140345705) derived		0.001				
T1095S (rs576112196) derived		0.001				
Q1104Stop (rs114751362) derived		0.001				
S1107N (rs71599965) derived	0.003	0.002	0.003	0.01		0.001
I1108V (rs529369991) derived						0.001
F1121I (rs541159378) derived		0.001				
V1134L (rs200396980) derived	0.002					0.01
D1136H (rs142122738) derived	0.001	0.002	0.001			
MGP (over 55 missenses)	Allele frequency of the variant					
	All	AFR	AMR	EUR	EAS	SAS
A8V (rs142330429) derived	0.02			0.004	0.031	0.043
A12T (rs184536928) derived			0.001			
G30S (rs531889747) derived						0.001
H48N (rs550400563) derived					0.001	
N64H (rs145210821) ancestral	0.002	0.001	0.01	0.001		
A77S (rs536781673) derived		0.002				
K78E (rs1801716) derived	0.004		0.004	0.003		0.014
R84Q (rs536440407) derived						0.001
E92K (rs550754490) derived				0.001		
N94H (rs539042522) derived		0.001				
R119H (rs546998085) derived		0.001				
R125Q (rs145254250) derived	0.001	0.004				
T127A (rs4236) ancestral	0.39	0.55	0.4	0.38	0.13	0.43



Frequency of archaic missense variants in living humans

