Association of genetic variants in enamel-formation genes with dental caries: A meta- and gene-cluster analysis

Xueyan Li1 PhD, Di Liu2 MD, Yang Sun3 PhD, Jingyun Yang4,5,6,7 PhD, Youcheng Yu3 PhD

1Department of Stomatology, Eye & Ent Hospital of Fudan University, Shanghai 200031, China
2Beijing Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing 100069, China
3Department of Stomatology, Zhongshan Hospital, Fudan University, Shanghai 200032, China
4Division of Statistics, School of Economics, Shanghai University, Shanghai 200444, China
5Research Center of Financial Information, Shanghai University, Shanghai 200444, China
6Rush Alzheimer’s Disease Center, Rush University Medical Center, Chicago, IL 60612, USA
7Department of Neurological Sciences, Rush University Medical Center, Chicago, IL 60612, USA

Correspondence should be addressed to
Youcheng Yu
Department of Stomatology
Zhongshan Hospital, Fudan University
180 Fengling Road
Shanghai, 200032
China
E-mail: yu.youcheng@zs-hospital.sh.cn

or

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Jingyun Yang
Division of Statistics
School of Economics
Shanghai University
99 Shangda Rd, Baoshan Dist
Shanghai, 200444
China
Tel: 86-21-66136063
Fax: 86-21-66134978
Email: jingyuny@gmail.com

Running title: Enamel formation genes and dental caries

Keywords: Enamel; dental caries; genetic variant; meta-analysis
Abstract

Previous studies have reported the association between multiple genetic variants in enamel formation-related genes and the risk of dental caries with inconsistent results. We performed a systematic literature search of the PubMed, Cochrane Library, HuGE and Google Scholar databases for studies published before March 21, 2020 and conducted meta-, gene-based and gene-cluster analysis on the association between genetic variants in enamel-formation-related genes and the risk of dental caries. Our systematic literature search identified 21 relevant publications including a total of 24 studies for analysis. The genetic variant rs17878486 in AMELX was significantly associated with dental caries risk (OR=1.40, 95% CI: 1.02-1.93, P=0.037). We found no significant association between the risk of dental caries with rs12640848 in ENAM (OR=1.15, 95% CI: 0.88-1.52, P=0.310), rs1784418 in MMP20 (OR=1.07, 95% CI: 0.76-1.49, P=0.702) and rs3796704 in ENAM (OR=1.06, 95% CI: 0.96-1.17, P=0.228). Gene-based analysis indicated that multiple genetic variants in AMELX showed joint association with the risk of dental caries (6 variants; P<10^-5), so did genetic variants in MMP13 (3 variants; P=0.004), MMP2 (3 variants; P<10^-5), MMP20 (2 variants; P<10^-5) and MMP3 (2 variants; P<10^-5). The gene-cluster analysis indicated a significant association between the genetic variants in this enamel-formation gene cluster and the risk of dental caries (P<10^-5). The present meta-analysis revealed that genetic variant rs17878486 in AMELX were associated with dental caries, and multiple genetic variants in enamel-formation-related genes jointly contribute to the risk of dental caries, supporting the role of genetic variants in the enamel-formation genes in the etiology of dental caries.
**Introduction**

Dental caries is one of the most common oral diseases, with an age-standardized global prevalence of untreated dentine carious lesions being around 9% in the primary dentition and around 35% in the permanent dentition during the last three decades\(^1\). Dental caries is a major public health concern, leading to tooth pain or loss and many other concomitants such as trouble in learning, eating or sleeping\(^2\). Dental caries remains to be very common despite the adoption of various preventive measures.

Dental caries results from continued localized demineralization of the dental enamel and dentine. It is a chronic disease with a multi-factorial etiology, involving the complex interactions between genetic, environmental and behavioral factors such as fluoride exposure, diet and oral hygiene\(^3-5\). Although previous studies have been successful in revealing the factors associated with the risk of dental caries\(^6-8\), more studies are needed to validate these findings, identify additional factors, and elucidate their exact roles in the etiology of dental caries.

The quality and quantity of enamel plays a direct role in the susceptibility to dental caries. Enamel formation related genes, such as AMBN, AMELX, TUFT1, KLK4 and ENAM, represent a cluster of genes that are involved in the pathway of odontogenesis of dentin-containing teeth\(^3,9\). Previous studies examined the association of genetic variants in this gene cluster with dental caries susceptibility, with inconsistent results\(^10-12\). Therefore, we performed this meta-analysis to examine the association between multiple genetic variants in this gene cluster and the risk of dental caries. Considering that the effect of individual genetic variant may be small, we also performed a gene-based and gene-
cluster analysis to explore the joint association of multiple genetic variants in this gene cluster with the risk of dental caries.

Materials and Methods

Ethical approval and informed consent statements are not required due to the systematic review and meta-analytic nature of this study.

Eligibility Criteria

We adopted the following inclusion criteria to determine study eligibility: 1) studies on human subjects; 2) the studies has case and control group, with the case group including subjects who had caries or high caries and the control group being care-free or having low/very low caries; and 3) the studies reported data on genetic variants in enamel formation genes for subjects in both the case group and the control group. We chose the studies with a larger sample size if multiple studies used overlapping data.

Two authors (XL and JY) performed an extensive literature search of the PubMed, Cochrane Library, HuGE and Google Scholar databases for studies published before March 21, 2020. The keywords used in the literature search are provided in the online supplementary file.

We retrieved all potentially relevant publications to evaluate study eligibility. We manually searched the references in all identified studies for research that might have been missed during the literature search. We also relied on Google Scholar’s ‘cited by’ tool to search for potential publications that cited the studies identified in the literature search. The two authors performed the literature search independently. The search was...
limited to studies published in English. Any disagreement was resolved by group discussion (XL, SY and JY).

**Data Extraction**

Two authors (DL and JY) independently extracted the following data from the eligible studies according to a pre-specified protocol for data extraction: name of the first author, year of publication, characteristics of the study participants, including sample size, mean age, distribution of gender, race/country of origin of the participants, screening method for dental caries, and genotype data for participants in the case and the control group. Any discrepancies were resolved in a group meeting. The quality of the included studies was assessed by two authors (DL and JY) independently using the Newcastle–Ottawa Scale (NOS)\(^\text{13}\). Extracted data were entered into a computerised spreadsheet for analysis.

**Data Analysis**

We used odds ratios (ORs) as a measure of the association between genetic variants in enamel-formation genes and the risk of dental caries. In all the meta-analyses, we used random-effects models to calculate the ORs and the corresponding 95% confidence intervals (CIs). The analyses were performed using different genetic models, including additive, allelic, dominant, recessive and co-dominant genetic models. Between-study heterogeneity was assessed using \(I^2\), and publication bias was visually checked by a contour-enhanced funnel plot and evaluated by Egger’s test. We performed meta-analysis for a single genetic variant when there were data from multiple studies for that genetic variant. However, we only reported meta-analysis results for genetic variants that had data from at least four studies. Meta-analysis results for genetic variants that had data
from less than four studies were used only in the gene-based and gene-cluster analysis as described in detail below.

**Gene-based and Gene-cluster Analysis**

We conducted gene-based analyses to assess the overall association of multiple genetic variants in each enamel-formation gene with the risk of dental caries risk. We followed the methods in Dr. Li et al. 2020\(^\text{14}\). Specifically, we used the P-values of all genetic variants within that gene obtained from our meta-analyses or from published literature when no meta-analysis could be done. Four different P-value combination methods were utilized—the Fisher’s method\(^\text{15}\), the Simes method\(^\text{16}\), the modified inverse normal method\(^\text{17}\) and the modified truncated product method (TPM)\(^\text{18, 19}\). For the modified TPM, we calculated unweighted and weighted TPM, where the former did not consider the difference in sample sizes whereas the latter employs the sample sizes as its weight, thereby allowing studies with larger sample sizes to play a larger role in the calculation\(^\text{18}\). A detailed description of the four methods has been given elsewhere\(^\text{19}\). To estimate the P-value for the modified TPM, we ran 100,000 simulations to account for the correlation between the P-values. Gene-cluster analysis followed a similar approach as gene-based analyses.

**Sensitivity/Additional Analysis**

We examined the association by including only the studies in which the genetic data in the control group satisfied Hardy–Weinberg equilibrium (HWE). We also repeated the meta-analyses by excluding studies of low quality (NOS < six stars). And finally, we
examined the association through meta-analysis by including only the studies that used data from children.

All statistical analyses were performed using R (https://www.r-project.org). A P-value <0.05 was considered statistically significant. This study was reported according to the PRISMA guidelines\textsuperscript{20}.

Results

Study Selection and Characteristics

Figure 1 shows the selection of eligible studies included in our meta-analyses. We identified 132 potential publications through our initial search. After screening the abstracts, 95 publications were excluded because they were not about human subjects, were not in English, were reviews/meta-analysis or were irrelevant. This left 37 studies that were retrieved for more detailed evaluation. We excluded an additional 16 studies because they were reviews or meta-analyses, the outcomes did not include dental caries, or because there were insufficient data. This resulted in 21 publications including a total of 24 studies that met the eligibility criteria and were included in our analyses\textsuperscript{5, 10-12, 21-38}.

In summary, the meta-analyses of rs12640848 in \textit{ENAM} included seven studies with a total of 1,256 subjects in the case group and 710 subjects in the control group\textsuperscript{12, 25, 27, 33, 34, 37}; the meta-analyses of rs1784418 in \textit{MMP20} included five studies with a total of 699 subjects in the case group and 817 subjects in the control group\textsuperscript{10, 24, 26, 30, 36}; the meta-analyses of rs17878486 in \textit{AMELX} included four studies with a total of 249 subjects in the case group and 193 subjects in the control group\textsuperscript{10, 12, 22, 25}; and the meta-analyses of rs3796704 in \textit{ENAM} included four studies with a total of 574 subjects in the case group.
and 533 subjects in the control group\textsuperscript{12, 25, 32, 34}. Data for other genetic variants came from fewer studies.

All included publications were published since 2008, with a sample size ranging from 71 to 1,005. The basic characteristics of the included studies are presented in Table 1. The majority of the included studies were of good quality, except four studies which had NOS scores of 4 or 5\textsuperscript{5, 22, 37, 38}.

Assessment of Publication Bias

We did not find evidence of a significant publication bias for the meta-analysis of rs12640848 in \textit{ENAM} (\(P=0.053\)), rs1784418 in \textit{MMP20} (\(P=0.238\)), rs17878486 in \textit{AMELX} (\(P=0.521\)) and rs3796704 in \textit{ENAM} (\(P=0.194\); Figure 2). Assessment of publication bias for the meta-analysis of other genetic variants is not very meaningful due to the limited number of studies included in the corresponding meta-analysis.

Association with the Risk of Dental Caries

For simplicity, we mainly reported results assuming an additive model for the meta-analyses. Complete results for meta-analyses and association analyses of individual genetic variant assuming different genetic models were presented in Supplementary Table 1-5. The genetic variant rs17878486 in \textit{AMELX} was significantly associated with dental caries risk (OR=1.40, 95\% CI: 1.02-1.93, \(P=0.037\); Table 2). We found that rs12640848 in \textit{ENAM}, the genetic variant that has largest number of studies in our meta-analysis, was not significantly associated with the risk of dental caries (OR=1.15, 95\% CI: 0.88-1.52, \(P=0.310\)). Meta-analysis also revealed no significant association of rs1784418 (OR=1.07, 95\% CI: 0.76-1.49, \(P=0.702\)) in \textit{MMP20} and rs3796704 (OR=1.06, 95\% CI: 0.96-1.17, \(P=0.228\)) in \textit{ENAM} with the risk of dental caries. There was significant
heterogeneity in the all the meta-analyses except the meta-analysis of rs3796704 (P=0.091; Figure 3). We did not find significant association of the four genetic variants with the risk of dental caries assuming other genetic models (Supplementary Table 2-5).

Meta-analyses of other genetic variants that included fewer studies revealed no significant association with the risk of dental caries. However, multiple genetic variants in ENAM, TUFT1, MMP2, MMP3, MMP8, MMP13, MMP20 and AMELX showed significant association with the risk of dental caries in individual studies (Supplementary Table 1).

Gene-based and Gene-cluster Analysis

Gene-based analysis results are presented in Table 3. We found that multiple genetic variants in AMELX showed joint association with the risk of dental caries (6 variants; all Ps<2×10⁻⁴), so did genetic variants in MMP13 (3 variants; all Ps≤0.01), MMP2 (3 variants; all Ps<10⁻⁵), MMP20 (2 variants; all Ps<10⁻⁵) and MMP3 (2 variants; all Ps<10⁻⁵). The gene-cluster analysis indicated a significant association between the genetic variants in this enamel-formation cluster and the risk of dental caries (all Ps<10⁻⁵; Table 4).

The x-axis is the odds ratio, and the y-axis is the standard error of the estimated effect on the risk of dental caries. The vertical line in the figure represents the overall estimated odds ratio. The two diagonal lines represent the pseudo 95% confidence limits of the effect estimate. Levels of statistical significance of the individual studies are indicated by the shaded regions: the white region corresponds to p-values greater than .05, the dark blue-shaded region corresponds to...
P-values between .025 and .05, the blue-shaded region corresponds to P-values between .01 and .025, and the light blue-shaded region corresponds to P-values below .01.

A) Funnel plot for meta-analysis of rs12640848; B) Funnel plot for meta-analysis of rs1784418; C) Funnel plot for meta-analysis of rs17878486; and D) Funnel plot for meta-analysis of rs3796704

Sensitivity Analysis
The association of rs17878486 in AMELX remained to be significant when we excluded studies in which genetic data in the control group violated HWE (OR=1.59, 95% CI: 1.15-2.20, P=0.006; Supplementary Table 6). The association disappeared after excluding studies of low quality and after excluding studies that used adult data (Supplementary Table 7-8). It should be noted that the total sample sizes for these sensitivity analyses were very limited. We did not find a significant association of the other three genetic variants with the risk of dental caries in all the sensitivity analyses (Supplementary Table 6-8).

Discussions
In this manuscript, we performed a systematic literature search and conducted meta-, gene-based and gene-cluster analysis to examine the association of multiple genetic variants in the enamel formation genes with the risk of dental caries. We found that rs17878486 in AMELX was significantly associated with the risk of dental caries, but no significant association of other genetic variants in the meta-analyses including at least four studies. However, gene-based analysis and gene-cluster analysis indicated that genetic variants in enamel-formation-related genes were jointly associated with the risk of dental caries. To the best of our knowledge, this is the first meta-analysis on some of
the genetic variants in the enamel-formation genes, and the first gene-based and gene-cluster analysis on the joint association of genetic variants in this gene cluster with the risk of dental caries.

AMELX is a gene located in both the X and Y chromosomes\(^{39}\) and encodes a set of isoforms of amelogenin, a major structural protein of the enamel organic matrix protein. Previous research using genetically engineered mice indicated AMELX was crucial for proper enamel formation\(^{40}\). Our meta-analysis of four studies including a total of 249 subjects with caries and 193 subjects without caries indicated that the genetic variant rs17878486 in AMELX was significantly associated with the risk of dental caries assuming an additive model; however, it showed no significant association under other genetic models (Supplementary Table 2-5). Moreover, no other genetic variants showed significant association except rs5933871 (Supplementary Table 1). Given the important role of AMELX in amelogenesis, future research is greatly needed to validate the relationship of the reported genetic variants and explore other genetic variants in this gene that may be associated with the risk of dental caries.

The gene ENAM is located on chromosome 4q 13.3 and has 10 exons. It encodes enamelin protein which is critical for the formation and elongation of enamel crystallites\(^{41}\). Meta-analysis for the genetic variant rs12640848 in ENAM has the largest number of included studies. Previous research suggested that this polymorphism was associated with the risk of dental caries\(^{42, 43}\). In our meta-analysis assuming an additive model, of the six publications including a total of seven studies, only one study including 541 caries-affected children and 177 caries-free children aged 13-15 years old in Czech indicated a significant positive association of the G allele with an increased risk of dental caries.
caries (OR=2.38, 95% CI: 2.12-2.67, \( P<0.0001 \))\textsuperscript{33}. However, we did notice that several included studies indicated a significant association of this variant under other genetic models\textsuperscript{27, 34}, although the overall meta-analysis results were still not significant (data not shown). Most of the other studied genetic variants in this gene were not associated with the risk of dental caries except one SNP rs3806804 which showed marginal association (OR=0.80, 95% CI: 0.65-1.00; \( P=0.049 \); Supplementary Table 1). More studies are needed to elucidate the role of rs12640848 and other genetic variants in ENAM in the formation of dental caries.

Matrix metalloproteinases (MMPs) refer to a large family of zinc-dependent endoproteinases which are fundamental in tooth formation and mineralization of dental tissue\textsuperscript{44}. More than 25 vertebrate MMPs have been identified, and 24 of them are present in humans\textsuperscript{45}. Genetic variants in many of the MMP genes, such as \textit{MMP2}, \textit{MMP3}, \textit{MMP13} and \textit{MMP20}, have been reported to be associated with the risk of dental caries\textsuperscript{23, 24, 28}. Meta-analysis of rs1784418 in \textit{MMP20} including a total of five studies showed no significant association of this genetic variant with the risk of dental caries (Table 1). However, many other genetic variants in this gene cluster were significantly associated with the risk of dental caries (Supplementary Table 1). These results should be interpreted with caution because they were based on data from single studies with limited sample sizes. Gene-based analysis indicated that genetic variants within \textit{MMP2}, \textit{MMP3}, \textit{MMP8}, \textit{MMP13} and \textit{MMP20} jointly associated with the risk of dental caries, further supporting the involvement of genetic variants in these genes in influencing the risk of dental caries. Future studies are warranted to reveal the function of these genetic variants in the etiology of dental caries.
We observed significant heterogeneities in all of the meta-analyses except the meta-
analysis for rs3796704 in ENAM (P=0.091; Table 2). The heterogeneity disappeared in
some sensitivity analyses by removing the studies that violated HWE in the control group,
the studies of low quality and the studies that used adult data, indicating that HWE, study
quality which included several items such as selection of the study subjects, and age of
study participants might be potential sources of heterogeneity. However, because of the
limited availability of data from each individual study, the exact source of the
heterogeneity could not track, and meta-regression is also not feasible and/or meaningful,
again due to the limited number of studies.

For the gene-based and gene-cluster analyses, we adopted the same approach as in our
previous work which revealed that, despite weak evidence of individual genetic variants
in actotransferrin (LTF), multiple genetic variants in LTF showed joint contribution to the
risk of dental caries\textsuperscript{14}. Both of our studies indicated that although the effect of a single
genetic variant may be small or insignificant, with proper methods we could still capture
the joint contribution of multiple genetic variants. However, our previous study was
limited to the exploration of genetic variants within a single gene, whereas in the present
study, we examined multiple genetic variants in multiple genes. As a result, we were able
to examine a larger number of genetic variants from a larger number of publications, and
findings from our gene-based and gene-cluster analysis represented the joint effects from
a larger of genetic variants.

Our study has some limitations. Despite efforts in the systematic literature search, the
sample sizes for many meta-analyses were limited. As a result, our findings need to be
validated by future studies with larger sample sizes. The association of the risk of dental
caries with many genetic variants was based on individual studies, which may be subject
to biases due to a number of factors such as small sample size and the genetic background
of the study participants. The exact relationship of these genetic variants with the risk of
dental caries warrant further research. We did not perform analysis by type of dentition
because of limited information and limited number of studies. Finally, due to a lack of
data for individual subjects, our meta-analyses did not control for important factors that
may affect the risk of dental caries. The estimated effect of the reported genetic variants
on the risk of dental caries could be greatly confounded by such factors, a limitation of
any meta-analysis that uses unadjusted analysis. Future research on the relationship
between enamel-formation-related genes and the risk of dental caries should take into
account the important confounding factors.

Conclusions

The present meta-analysis revealed that genetic variant rs17878486 in *AMELX* was
associated with dental caries, and multiple genetic variants in enamel-formation-related
genes jointly contribute to the risk of dental caries. Future studies with large sample sizes
that control for important confounding factors, such as diet, microbial and host
characteristics, are needed to validate our findings and to explore additional genetic loci
in this gene cluster that might also affect the risk of dental caries.

Data Availability

No additional data are available.
Conflict of Interest

We have no conflicts of interest to disclose.

Funding Statement

This study was supported by the National Natural Science Foundation of China (Grant No. 81771493). Dr. Jingyun Yang’s research was also supported by NIH/NIA grant R01AG036042 and the Illinois Department of Public Health. Di Liu’s research was supported by China Scholarship Council (CSC 201908110339). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Supplementary Materials

Supplementary Table 1-5 provided results for the association of genetic variants in the enamel-formation genes with the risk of dental caries assuming different genetic models. Supplementary Table 6-8 provided results for sensitivity/subgroup analysis in meta-analyzing four genetic variants in the enamel-formation genes in association with the risk of dental caries.
References


Figure legends

**Figure 1.** Flow diagram of the process of selecting studies included in the meta-analyses.

Note: Please see the Methods section for additional details.

**Figure 2.** Contour-enhanced funnel plots for meta-analyses of the association with dental caries assuming an additive model.

**Figure 3.** Forest plots for meta-analysis of the association with dental caries assuming an additive model.

Each study is represented by a square whose area is proportional to the weight of the study. The overall effect from meta-analysis is represented by a diamond whose width represents the 95% CI for the estimated OR.

A) Forest plot for meta-analysis of rs12640848; B) Forest plot for meta-analysis of rs1784418; C) Forest plot for meta-analysis of rs17878486; and D) Forest plot for meta-analysis of rs3796704

OR, odds ratio; CI, confidence interval.
### Table 1. Basic characteristics of all the studies included in the analyses.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year of publication</th>
<th>Country/origin</th>
<th>Ethnicity</th>
<th>Cases</th>
<th>Male (%)</th>
<th>Control</th>
<th>Diagnosis of dental caries</th>
<th>NOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ouryouji et al., 2008</td>
<td>2008</td>
<td>India</td>
<td>Asian</td>
<td>80</td>
<td>5.4±0.92</td>
<td>63.8%</td>
<td>67</td>
<td>4.8±1.2</td>
</tr>
<tr>
<td>Kang et al., 2011</td>
<td>2011</td>
<td>Korea</td>
<td>Asian</td>
<td>87</td>
<td>23.2±7.8</td>
<td>73.4%</td>
<td>33</td>
<td>21.3±7.8</td>
</tr>
<tr>
<td>Olszowski et al., 2012-a</td>
<td>2012</td>
<td>Poland</td>
<td>European</td>
<td>37</td>
<td>5</td>
<td>51.4%</td>
<td>34</td>
<td>5</td>
</tr>
<tr>
<td>Olszowski et al., 2012-b</td>
<td>2012</td>
<td>Poland</td>
<td>European</td>
<td>58</td>
<td>13</td>
<td>31.0%</td>
<td>50</td>
<td>13</td>
</tr>
<tr>
<td>Tannure et al., 2012-a</td>
<td>2012</td>
<td>Brazil</td>
<td>Mixed</td>
<td>293</td>
<td>8.83±2.95</td>
<td>51.9%</td>
<td>212</td>
<td>9.40±3.39</td>
</tr>
<tr>
<td>Tannure et al., 2012-b</td>
<td>2012</td>
<td>Brazil</td>
<td>Mixed</td>
<td>227</td>
<td>8.73±0.17</td>
<td>50.2%</td>
<td>161</td>
<td>9.45±0.22</td>
</tr>
<tr>
<td>Ergöz et al., 2013</td>
<td>2013</td>
<td>Turkey</td>
<td>European</td>
<td>100</td>
<td>8.65±1.94</td>
<td>50%</td>
<td>100</td>
<td>8.65±1.94</td>
</tr>
<tr>
<td>Jeremias et al., 2013#</td>
<td>2013</td>
<td>Brazil</td>
<td>Mixed</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Yıldız et al., 2015</td>
<td>2015</td>
<td>Turkey</td>
<td>Mixed</td>
<td>77</td>
<td>20-60</td>
<td>-</td>
<td>77</td>
<td>20-60</td>
</tr>
<tr>
<td>Antunes et al., 2016</td>
<td>2016</td>
<td>Brazil</td>
<td>Mixed</td>
<td>245</td>
<td>4.38±1.18</td>
<td>53%</td>
<td>541</td>
<td>3.09±1.49</td>
</tr>
<tr>
<td>Gerreth et al., 2016</td>
<td>2016</td>
<td>Poland</td>
<td>European</td>
<td>48</td>
<td>20-42 months</td>
<td>47.9%</td>
<td>48</td>
<td>20-42 months</td>
</tr>
<tr>
<td>Karayasheva et al., 2016</td>
<td>2016</td>
<td>Bulgaria</td>
<td>European</td>
<td>82</td>
<td>20-32</td>
<td>-</td>
<td>20</td>
<td>20-32</td>
</tr>
<tr>
<td>Cavallari et al., 2017</td>
<td>2017</td>
<td>Brazil</td>
<td>Mixed</td>
<td>100</td>
<td>18.43</td>
<td>51.3%</td>
<td>100</td>
<td>18.52</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Country</td>
<td>Region</td>
<td>Sample Size</td>
<td>Age Range</td>
<td>Age Mean±SD</td>
<td>Caries Prevalence</td>
<td>Missing Teeth</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------</td>
<td>---------</td>
<td>--------</td>
<td>-------------</td>
<td>------------</td>
<td>-------------</td>
<td>------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Filho et al., 2017</td>
<td>2017</td>
<td>Brazil</td>
<td>Mixed</td>
<td>103</td>
<td>4-7</td>
<td>49.5%</td>
<td>81</td>
<td>4-7</td>
</tr>
<tr>
<td>Gerreth et al., 2017</td>
<td>2017</td>
<td>Poland</td>
<td>European</td>
<td>48</td>
<td>20-42</td>
<td>47.9%</td>
<td>48</td>
<td>20-42</td>
</tr>
<tr>
<td>Wang et al., 2017</td>
<td>2017</td>
<td>China</td>
<td>Asian</td>
<td>505</td>
<td>10-15</td>
<td>41.7±7.0</td>
<td>52.7%</td>
<td>500</td>
</tr>
<tr>
<td>Koohpeima et al., 2018</td>
<td>2018</td>
<td>Iran</td>
<td>Asian</td>
<td>236</td>
<td>20-42</td>
<td>29.8±7.9</td>
<td>37.3%</td>
<td>166</td>
</tr>
<tr>
<td>Linhartova et al., 2017 -a</td>
<td>2018</td>
<td>Czech</td>
<td>European</td>
<td>109</td>
<td>2-6</td>
<td>58.7%</td>
<td>78</td>
<td>2-6</td>
</tr>
<tr>
<td>Linhartova et al., 2017 -b</td>
<td>2018</td>
<td>Czech</td>
<td>European</td>
<td>541</td>
<td>13-15</td>
<td>52.7%</td>
<td>177</td>
<td>13-15</td>
</tr>
<tr>
<td>Divakar et al., 2019</td>
<td>2019</td>
<td>India</td>
<td>Asian</td>
<td>168</td>
<td>12-15</td>
<td>6.9±1.9</td>
<td>52.4%</td>
<td>193</td>
</tr>
<tr>
<td>Hu et al., 2019</td>
<td>2019</td>
<td>China</td>
<td>Asian</td>
<td>161</td>
<td>12-15</td>
<td>52.2%</td>
<td>196</td>
<td>12-15</td>
</tr>
<tr>
<td>Vasconcelos et al., 2019</td>
<td>2019</td>
<td>Brazil</td>
<td>Mixed</td>
<td>131</td>
<td>10-12</td>
<td>-</td>
<td>85</td>
<td>10-12</td>
</tr>
<tr>
<td>Duran-Merino et al., 2020</td>
<td>2020</td>
<td>Mexico</td>
<td>Mixed</td>
<td>39</td>
<td>11</td>
<td>-</td>
<td>32</td>
<td>11</td>
</tr>
<tr>
<td>Nicoline et al., 2020</td>
<td>2020</td>
<td>Indonesia</td>
<td>Asian</td>
<td>95</td>
<td>-</td>
<td>-</td>
<td>89</td>
<td>-</td>
</tr>
</tbody>
</table>

Data for age were presented as mean, mean±SD or range.

*Teeth were evaluated by one trained and calibrated dentist specialized in pediatric dentistry using an artificial light, a dental mirror and a probe, after calibration by an experienced specialist.

#Number of subjects varies for different genetic variants

SD: standard deviation; NOS: the Newcastle–Ottawa scale; DMFT: decayed, missing and filled teeth index; DMFS, decayed, missing, and filled surfaces; ICDAS, International Decay Detection and Assessment System.
Table 2. Meta-analysis of the association of genetic variants in enamel-formation genes with the risk of dental caries.*

<table>
<thead>
<tr>
<th>Genetic variant*</th>
<th>Gene</th>
<th>Number of studies</th>
<th>Number of study subjects</th>
<th>Test of heterogeneity</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Dental caries</td>
<td>Control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12640848</td>
<td>ENAM</td>
<td>7</td>
<td>1,256</td>
<td>7.69×10^{-25}</td>
<td>1.15 (0.88 - 1.52)</td>
<td>0.310</td>
</tr>
<tr>
<td>rs1784418</td>
<td>MMP20</td>
<td>5</td>
<td>699</td>
<td>1.38×10^{-20}</td>
<td>1.07 (0.76 - 1.49)</td>
<td>0.702</td>
</tr>
<tr>
<td>rs17878486</td>
<td>AMELX</td>
<td>4</td>
<td>249</td>
<td>3.58×10^{-4}</td>
<td>1.40 (1.02 - 1.93)</td>
<td>0.037</td>
</tr>
<tr>
<td>rs3796704</td>
<td>ENAM</td>
<td>4</td>
<td>574</td>
<td>0.091</td>
<td>1.06 (0.96 - 1.17)</td>
<td>0.228</td>
</tr>
</tbody>
</table>

*Assuming an additive genetic model

We only reported results for genetic variants that had data from at least four studies. Results for other genetic variants assuming different genetic models were reported in Supplementary Table 1-5.

OR: odds ratio; CI: confidence interval.
Table 3. Gene-based analysis of the association of genetic variants in enamel-formation genes with the risk of dental caries.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Number of variants</th>
<th>Fisher</th>
<th>Simes</th>
<th>Inverse</th>
<th>TPM (unweighted)</th>
<th>TPM (weighted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENAM</td>
<td>6</td>
<td>0.395</td>
<td>0.296</td>
<td>0.485</td>
<td>0.305</td>
<td>0.324</td>
</tr>
<tr>
<td>AMELX</td>
<td>6</td>
<td>1.26×10^-4</td>
<td>2.98×10^-5</td>
<td>1.40×10^-5</td>
<td>&lt;10^-5</td>
<td>&lt;10^-5</td>
</tr>
<tr>
<td>KLK4</td>
<td>5</td>
<td>0.494</td>
<td>0.276</td>
<td>0.599</td>
<td>0.275</td>
<td>0.381</td>
</tr>
<tr>
<td>TUFT1</td>
<td>4</td>
<td>0.103</td>
<td>0.077</td>
<td>0.064</td>
<td>0.090</td>
<td>0.088</td>
</tr>
<tr>
<td>MMP13</td>
<td>3</td>
<td>0.006</td>
<td>0.002</td>
<td>0.010</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>AMBN</td>
<td>3</td>
<td>0.452</td>
<td>0.262</td>
<td>0.998</td>
<td>0.261</td>
<td>0.199</td>
</tr>
<tr>
<td>MMP2</td>
<td>3</td>
<td>&lt;10^-5</td>
<td>4.60×10^-18</td>
<td>2.00×10^-5</td>
<td>&lt;10^-5</td>
<td>&lt;10^-5</td>
</tr>
<tr>
<td>MMP11</td>
<td>2</td>
<td>0.514</td>
<td>0.605</td>
<td>0.459</td>
<td>0.149</td>
<td>0.150</td>
</tr>
<tr>
<td>MMP20</td>
<td>2</td>
<td>&lt;10^-5</td>
<td>3.48×10^-19</td>
<td>8.73×10^-18</td>
<td>&lt;10^-5</td>
<td>&lt;10^-5</td>
</tr>
<tr>
<td>MMP8</td>
<td>2</td>
<td>1.71×10^-4</td>
<td>0.005</td>
<td>0.004</td>
<td>0.004</td>
<td>0.003</td>
</tr>
<tr>
<td>MMP3</td>
<td>2</td>
<td>1.51×10^-13</td>
<td>1.59×10^-8</td>
<td>7.55×10^-9</td>
<td>&lt;10^-5</td>
<td>&lt;10^-5</td>
</tr>
<tr>
<td>MMP9</td>
<td>1</td>
<td>0.705</td>
<td>0.705</td>
<td>0.705</td>
<td>0.705</td>
<td>0.705</td>
</tr>
</tbody>
</table>

TPM, the modified truncated product method
Table 4. Gene-cluster analysis of the association of genetic variants in the enamel-formation genes with the risk of dental caries.

<table>
<thead>
<tr>
<th>Gene cluster</th>
<th>Fisher</th>
<th>Simes</th>
<th>Inverse</th>
<th>TPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enamel-formation genes</td>
<td>$&lt;10^{-5}$</td>
<td>$4.17 \times 10^{-18}$</td>
<td>$2.77 \times 10^{-45}$</td>
<td>$&lt;10^{-5}$</td>
</tr>
</tbody>
</table>

TPM, the modified truncated product method
Records identified through database searching (n = 162)

Additional records identified through other sources (n = 2)

Records after duplicates removed (n = 132)

Records excluded (n = 95)
- Not in English
- Not human studies
- Review/meta-analysis
- Irrelevant

Records screened (n = 132)

Full-text articles assessed for eligibility (n = 37)

Full-text articles excluded, with reasons (n = 16)
- Review/meta-analysis
- Outcome not including dental caries
- No sufficient data

Studies included in qualitative synthesis (n = 21)

Studies included in quantitative synthesis (n = 21)
### A) Odds Ratio

<table>
<thead>
<tr>
<th>Study</th>
<th>TE</th>
<th>SE</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ergöz et al., 2013</td>
<td>-0.04</td>
<td>0.0852</td>
<td>15.5%</td>
<td>0.96 [0.81; 1.13]</td>
</tr>
<tr>
<td>Jeremias et al., 2013</td>
<td>0.02</td>
<td>0.1278</td>
<td>14.4%</td>
<td>1.02 [0.80; 1.30]</td>
</tr>
<tr>
<td>Gerreth et al., 2016</td>
<td>-0.26</td>
<td>0.1894</td>
<td>10.2%</td>
<td>0.77 [0.56; 1.07]</td>
</tr>
<tr>
<td>Linhartova et al., 2018-a</td>
<td>0.21</td>
<td>0.1028</td>
<td>15.1%</td>
<td>1.23 [1.00; 1.50]</td>
</tr>
<tr>
<td>Linhartova et al., 2018-b</td>
<td>0.87</td>
<td>0.0589</td>
<td>16.0%</td>
<td>2.38 [2.12; 2.67]</td>
</tr>
<tr>
<td>Divakar et al., 2019</td>
<td>0.11</td>
<td>0.1281</td>
<td>14.4%</td>
<td>1.12 [0.87; 1.44]</td>
</tr>
<tr>
<td>Duran-Merino et al., 2020</td>
<td>-0.05</td>
<td>0.2296</td>
<td>11.4%</td>
<td>0.95 [0.61; 1.49]</td>
</tr>
</tbody>
</table>

**Total (95% CI):** 100.0%  **1.15 [0.88; 1.52]**

Heterogeneity: Tau² = 0.1193; Chi² = 126.31, df = 6 (P < 0.01); I² = 95%

### B) Odds Ratio

<table>
<thead>
<tr>
<th>Study</th>
<th>TE</th>
<th>SE</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tannure et al., 2012-b</td>
<td>0.24</td>
<td>0.0766</td>
<td>20.8%</td>
<td>1.27 [1.09; 1.47]</td>
</tr>
<tr>
<td>Antunes et al., 2016</td>
<td>-0.56</td>
<td>0.0639</td>
<td>21.0%</td>
<td>0.57 [0.50; 0.64]</td>
</tr>
<tr>
<td>Filho et al., 2017</td>
<td>0.34</td>
<td>0.1232</td>
<td>19.5%</td>
<td>1.40 [1.10; 1.78]</td>
</tr>
<tr>
<td>Gerreth et al., 2017</td>
<td>0.03</td>
<td>0.1424</td>
<td>18.8%</td>
<td>1.03 [0.78; 1.36]</td>
</tr>
<tr>
<td>Vasconcelos et al., 2019</td>
<td>0.32</td>
<td>0.1110</td>
<td>19.9%</td>
<td>1.38 [1.11; 1.71]</td>
</tr>
</tbody>
</table>

**Total (95% CI):** 100.0%  **1.07 [0.76; 1.49]**

Heterogeneity: Tau² = 0.1335; Chi² = 99.31, df = 4 (P < 0.01); I² = 96%

### C) Odds Ratio

<table>
<thead>
<tr>
<th>Study</th>
<th>TE</th>
<th>SE</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kang et al., 2011</td>
<td>0.51</td>
<td>0.1071</td>
<td>27.6%</td>
<td>1.67 [1.36; 2.06]</td>
</tr>
<tr>
<td>Ergöz et al., 2013</td>
<td>0.14</td>
<td>0.1789</td>
<td>22.7%</td>
<td>1.15 [0.81; 1.64]</td>
</tr>
<tr>
<td>Jeremias et al., 2013</td>
<td>0.02</td>
<td>0.0916</td>
<td>28.5%</td>
<td>1.02 [0.86; 1.23]</td>
</tr>
<tr>
<td>Gerreth et al., 2017</td>
<td>0.74</td>
<td>0.1992</td>
<td>21.3%</td>
<td>2.10 [1.42; 3.10]</td>
</tr>
</tbody>
</table>

**Total (95% CI):** 100.0%  **1.40 [1.02; 1.93]**

Heterogeneity: Tau² = 0.0841; Chi² = 18.43, df = 3 (P < 0.01); I² = 84%

### D) Odds Ratio

<table>
<thead>
<tr>
<th>Study</th>
<th>TE</th>
<th>SE</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ergöz et al., 2013</td>
<td>0.02</td>
<td>0.0795</td>
<td>22.1%</td>
<td>1.02 [0.87; 1.19]</td>
</tr>
<tr>
<td>Jeremias et al., 2013</td>
<td>0.00</td>
<td>0.0894</td>
<td>19.2%</td>
<td>1.00 [0.84; 1.19]</td>
</tr>
<tr>
<td>Kooppeima et al., 2018</td>
<td>0.18</td>
<td>0.0535</td>
<td>32.1%</td>
<td>1.20 [1.08; 1.33]</td>
</tr>
<tr>
<td>Divakar et al., 2019</td>
<td>-0.01</td>
<td>0.0667</td>
<td>26.6%</td>
<td>0.99 [0.87; 1.13]</td>
</tr>
</tbody>
</table>

**Total (95% CI):** 100.0%  **1.06 [0.96; 1.17]**

Heterogeneity: Tau² = 0.0047; Chi² = 6.46, df = 3 (P = 0.09); I² = 54%