Enamel Formation Genes Are Associated with High Caries Experience in Turkish Children

A. Patir a  F. Seymen a  M. Yildirim a  K. Deeley b  M.E. Cooper b
M.L. Marazita b, d, e  A.R. Vieira b, c, d

a Department of Pedodontics, Istanbul University, Istanbul, Turkey; b Department of Oral Biology and Center for Craniofacial and Dental Genetics and c Department of Pediatric Dentistry, School of Dental Medicine, d Department of Human Genetics, Graduate School of Public Health, and e Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, Pa., USA

Key Words
Ameloblastin · Amelogenin · Enamelin · Tuftelin · Tuftelin interacting protein 11

Abstract
There is evidence for a genetic component in caries susceptibility, and studies in humans have suggested that variation in enamel formation genes and their interaction with Streptococcus mutans levels may contribute to caries. For the present study, we used DNA samples collected from 173 unrelated children from Istanbul: 91 children with 4 or more affected tooth surfaces and 82 caries-free children. Six single-nucleotide polymorphism markers were genotyped in selected candidate genes (ameloblastin, amelogenin, enamelin, tuftelin, tuftelin 1 and tuftelin interacting protein 11) that influence enamel formation. Allele and genotype frequencies were compared between groups with distinct caries experience. Regression analysis was used for the evaluation of individual gene effects, environmental effects and gene-environment interactions. Overrepresentation of the C allele of the amelogenin marker was seen in cases with dmft scores higher than 8 (p = 0.01) when compared to controls. Also, overrepresentation of the T allele of the ameloblastin marker was seen in cases with dmfs scores higher than 10 (p = 0.05) when compared to controls. In addition, the CT genotype of the tuftelin rs3790506 marker was overrepresented in cases with dmft scores higher than 5 (p = 0.05) and dmfs scores higher than 6 (p = 0.05) compared to controls. The best-fitting model showed that dmfs is increased when the following factors are present: (1) females and both the anterior and posterior teeth are affected simultaneously, (2) when the T allele of the tuftelin rs3790506 is involved, and (3) the C allele of the amelogenin rs17878486 is involved. Our study provides support that genes involved in enamel formation modify caries susceptibility in humans.

A genetic susceptibility to caries has been proposed based on twin studies [reviewed by Shuler, 2001], family studies [Klein and Palmer, 1940; Klein, 1946] and animal models [Hunt et al., 1944; Rosen et al., 1961]. The earlier molecular genetic studies in humans focused on proline-rich peptides in saliva [Yu et al., 1986; Ayad et al., 2000; Young et al., 2002; Zakhary et al., 2007] and enamel formation genes [Slayton et al., 2005; Deeley et al., 2008]. More recently, a genome-wide scan searching for caries susceptibility and protective loci was performed, and suggestive loci influencing caries were proposed [Vieira et al., 2008].
Amelogenesis is under genetic control, and the size, shape, shade and even caries susceptibility can be affected by genetic variation [Simmer and Hu, 2001]. Regarding enamel formation genes, tuftelin genotypes appeared to interact with levels of *Streptococcus mutans* infection in children with early childhood caries. In this study, the authors investigated children from 3 to 5 years of age, and cases were defined as having 4 or more decayed or filled tooth surfaces, while controls were defined as children with no evidence of caries (including white-spot lesions) [Slayton et al., 2005]. In adults, genetic variation in *amelogenin* was associated with higher caries experience [Deeley et al., 2008].

It is well documented that the first studies of complex traits suggest a stronger genetic effect than is found by subsequent studies. Both bias and genuine population diversity might explain why early studies tend to overestimate the disease predisposition conferred by candidate gene polymorphisms [Ioannidis et al., 2001]. Our hypothesis is that variation in genes involved in enamel formation contributes to increased caries susceptibility in humans. Therefore, we investigated genetic markers in genes involved in enamel formation that were previously studied [Slayton et al., 2005; Deeley et al., 2008] to provide independent replication of the original results, which suggested that variation in *tuftelin* in the presence of *S. mutans* and variation in *amelogenin* contributed to higher caries experience in humans.

### Subjects and Methods

#### Subject Screening and Sample Collection

Subjects were recruited for this study as approved by both the Istanbul University and University of Pittsburgh Institutional Review Boards, and informed consent was obtained. Eligible children were from 3 to 6 years of age and were enrolled in the Pediatric Dentistry Clinics of Istanbul University and daycare facilities in the city of Istanbul. One of the authors (A.P.) carried out the clinical examination after being calibrated by an experienced specialist (F.S.). The intraexaminer agreement was assessed by a second clinical examination in 10% of the sample after 2 weeks, with a κ of 1.0. Cases were defined as children with 4 or more decayed or filled tooth surfaces, while controls were defined as children with no evidence of caries (including no evidence of white-spot lesions) and no history of caries. Examinations were done with the use of a flashlight and mouth mirror. Caries experience was scored by the dmft and dmfs indexes according to World Health Organization guidelines. Subjects with dmfs $>$ 30 but $<$ 4 were considered intermediate in phenotype and were not included in the current study. Most parents reported not brushing the teeth of their children. Drinking water in the region is not artificially fluoridated. Unstimulated saliva samples were obtained from all participants (subjects were asked to spit) and stored in Oragene DNA Self-Collection kits (DNA Genotek Inc.) at room temperature until being processed. No centrifugation was performed in the saliva samples. No plaque samples were collected. DNA was extracted according to the manufacturer’s instructions.

*S. mutans* Infection Detection

A real-time polymerase chain reaction assay was used to determine the presence of *S. mutans* in the saliva of the subjects [Yano et al., 2002]. The assay implements primers targeting *glt* genes of *S. mutans*. We used the assay to test if detectable levels of *S. mutans* were present in the human genomic DNA samples we extracted from the saliva.

#### Genotyping and Statistical Analysis

Genotyping was carried out using Taqman chemistry [Ranade et al., 2001] and performed on an Applied Biosystems 7900 HT Sequence Detection System machine. Single-nucleotide polymorphism markers were selected within candidate genes involved in enamel formation (table 1).

$\chi^2$ and Student $t$ tests were performed to compare gender and age frequencies, and to test for deviations in the genotype distributions from Hardy-Weinberg equilibrium. In addition, allele frequency comparisons between cases and controls were performed taking into consideration severity of caries (higher vs. lower dmft and dmfs scores), presence of *S. mutans* and types of affected teeth (posterior vs. anterior). Because there was a statistically significant difference between the mean ages of the two comparison groups (table 2), we used regression analysis to account for the age difference and to investigate main effect models.
after which interactions (gene/environment, i.e. single-nucleotide polymorphism/S. mutans) with covariates-confounders (gender, teeth affected) were tested. A power analysis demonstrated that a statistically significant correlation of 0.3 – with a 2-sided test to 0.05 and a power of 80% – can be found with a sample of 78 or more.

### Results

All genotypes were in Hardy-Weinberg equilibrium. Tables 3 and 4 summarize the results of the allele frequency comparisons and regression analysis. Overrepresentation of the C allele of the amelogenin marker was

---

### Table 2. Demographics and caries prevalence of study subjects

<table>
<thead>
<tr>
<th></th>
<th>Cases (n = 91)</th>
<th>Controls (n = 82)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± standard deviation), years</td>
<td>4.82 ± 0.81</td>
<td>5.99 ± 0.11</td>
<td>t test, p &lt; 0.001</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>52</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>39</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>dmft (mean)</td>
<td>7.2 ± 2.3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>dmfs (mean)</td>
<td>11.8 ± 7.7</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Affected region</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior (molars) only</td>
<td>57</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Anterior and posterior</td>
<td>34</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Streptococcus mutans detected</td>
<td>Yes</td>
<td>68</td>
<td>17</td>
</tr>
<tr>
<td>No</td>
<td>23</td>
<td>65</td>
<td></td>
</tr>
</tbody>
</table>

#### Table 3. Summary of the allele and genotype frequency comparisons (dmft/dmfs p values)

<table>
<thead>
<tr>
<th>Marker</th>
<th>Cases vs. controls</th>
<th>Moderate + severe cases1 vs. controls</th>
<th>Mild cases2 vs. controls</th>
<th>Moderate + severe cases3 vs. controls</th>
<th>Severe cases4 vs. controls</th>
<th>Moderate + severe cases vs. mild cases (analysis within cases)</th>
<th>Mild vs. moderate severe cases (analysis within cases)</th>
<th>S. mutans</th>
<th>Cases with posterior teeth affected vs. cases with anterior + posterior teeth affected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs34538475 AMBN</td>
<td>0.1</td>
<td>0.06/0.07</td>
<td>1.0/0.7</td>
<td><strong>0.04/0.38</strong></td>
<td><strong>0.66/0.05</strong></td>
<td>0.36/0.7</td>
<td>0.45/0.64</td>
<td>0.47</td>
<td>1.0</td>
</tr>
<tr>
<td>rs17878486 AMELX</td>
<td>0.3</td>
<td>0.2/0.2</td>
<td>0.47/0.62</td>
<td><strong>0.01/0.09</strong></td>
<td><strong>0.31/0.37</strong></td>
<td><strong>0.03/0.27</strong></td>
<td><strong>0.52/0.05</strong></td>
<td></td>
<td><strong>0.008</strong></td>
</tr>
<tr>
<td>rs3796704 ENAM</td>
<td>0.11</td>
<td>0.14/0.2</td>
<td>0.28/0.24</td>
<td><strong>0.19/0.1</strong></td>
<td><strong>0.22/0.39</strong></td>
<td>0.53/0.22</td>
<td><strong>0.6/0.83</strong></td>
<td>0.27</td>
<td>1.0</td>
</tr>
<tr>
<td>rs2337360 TUFT1</td>
<td>0.2</td>
<td>0.2/0.23</td>
<td>0.64/0.61</td>
<td><strong>0.15/0.12</strong></td>
<td><strong>0.96/0.76</strong></td>
<td>0.71/0.67</td>
<td>0.24/0.23</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>rs3790506 TUFT1</td>
<td>0.71</td>
<td>0.48/0.5</td>
<td>0.24/0.78</td>
<td><strong>0.41/0.32</strong></td>
<td><strong>0.13/0.16</strong></td>
<td>0.21/0.53</td>
<td>0.84/0.52</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>rs134136 TFIP11</td>
<td>0.77</td>
<td>0.55/0.58</td>
<td>0.18/0.94</td>
<td><strong>0.55/0.69</strong></td>
<td><strong>0.76/0.80</strong></td>
<td><strong>0.11/0.24</strong></td>
<td><strong>0.27/0.94</strong></td>
<td>0.96</td>
<td>0.34</td>
</tr>
<tr>
<td>Genotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rs34538475 AMBN</td>
<td>0.26</td>
<td>0.1/0.14</td>
<td>1.0/0.15</td>
<td><strong>0.08/0.36</strong></td>
<td><strong>0.66/0.09</strong></td>
<td>0.38/0.71</td>
<td><strong>0.57/0.10</strong></td>
<td>0.55</td>
<td>1.0</td>
</tr>
<tr>
<td>rs17878486 AMELX</td>
<td>0.63</td>
<td>0.53/0.62</td>
<td>1.0/0.56</td>
<td><strong>0.62/0.63</strong></td>
<td><strong>0.1/0.33</strong></td>
<td>0.83/0.74</td>
<td><strong>0.21/0.18</strong></td>
<td>0.79</td>
<td><strong>0.05</strong></td>
</tr>
<tr>
<td>rs3796704 ENAM</td>
<td>0.11</td>
<td>0.21/0.32</td>
<td>0.28/0.19</td>
<td><strong>0.19/0.1</strong></td>
<td><strong>0.22/0.28</strong></td>
<td>0.53/0.22</td>
<td><strong>0.72/0.9</strong></td>
<td>0.26</td>
<td>1.0</td>
</tr>
<tr>
<td>rs2337360 TUFT1</td>
<td>0.19</td>
<td>0.23/0.26</td>
<td>0.88/0.25</td>
<td><strong>0.36/0.51</strong></td>
<td><strong>0.35/0.17</strong></td>
<td><strong>1.0/1.0</strong></td>
<td><strong>0.96/0.89</strong></td>
<td>0.42</td>
<td>0.35</td>
</tr>
<tr>
<td>rs3790506 TUFT1</td>
<td>0.1</td>
<td><strong>0.05/0.05</strong></td>
<td>0.26/0.56</td>
<td><strong>0.05/0.03</strong></td>
<td><strong>0.56/0.27</strong></td>
<td><strong>0.08/0.14</strong></td>
<td><strong>0.22/0.23</strong></td>
<td>0.89</td>
<td>0.64</td>
</tr>
<tr>
<td>rs134136 TFIP11</td>
<td>0.79</td>
<td>0.78/0.8</td>
<td>0.1/0.85</td>
<td><strong>0.8/0.8</strong></td>
<td><strong>1.0/0.97</strong></td>
<td><strong>0.11/0.21</strong></td>
<td><strong>0.25/0.98</strong></td>
<td>0.98</td>
<td>0.59</td>
</tr>
</tbody>
</table>

For explanation of gene symbols, see table 1. Bold indicates p values equal or below 0.05.
1 Cases with dmft scores higher than 4 or dmfs scores higher than 6.
2 Cases with dmft scores equal to 4 or dmfs scores up to 6.
3 Cases with dmft scores between 5 and 8 inclusive or dmfs scores between 6 and 10 inclusive.
4 Cases with dmft scores higher than 8 or dmfs scores higher than 10.
5 Significant result for the C allele in cases with both anterior and posterior teeth affected.
6 Significant result for the CC genotype in cases with both anterior and posterior teeth affected.
seen in cases with dmft scores higher than 8 (p = 0.01) when compared to controls, and when compared to cases with dmft scores between 5 and 8 (p = 0.03). In cases with both anterior and posterior teeth affected (p = 0.008), overrepresentation of the C allele of the amelogenin marker was seen when compared with cases with only posterior teeth affected. Also, overrepresentation of the T allele of the ameloblastin marker was seen in cases with dmft scores higher than 8 (p = 0.01) when compared to controls, and when compared to cases with dmft scores between 5 and 8 (p = 0.03). In cases with both anterior and posterior teeth affected (p = 0.008), overrepresentation of the C allele of the amelogenin marker was seen when compared with cases with only posterior teeth affected. For modeling dmfs (33 missing observations due to missing values of SNPs)

Base model: intercept
Null: Model: gender as a confounder $R^2 = 0.0$
Nullb: Model: age as confounder $R^2 = 0.312$ As age increases, dmfs decreases
Ia: Model: rs34538475 $R^2 = 0.027$ Presence of T allele increases dmfs
Ib: Model: rs17878486 $R^2 = 0.019$ Presence of C allele increases dmfs
IIa: Model: age + rs17878486 $R^2 = 0.331$***
IIb: Model: age + rs34538475 $R^2 = 0.317$
IIc: Model: (age + gender) + age $R^2 = 0.314$
III: Model: age + rs17878486 + rs34538475 $R^2 = 0.343$

Region = Posterior teeth affected versus anterior + posterior teeth affected; SNP = single-nucleotide polymorphism. The asterisks indicate best-fitting models.
the amelogenin marker was seen in cases with both anterior and posterior teeth affected \((p = 0.05)\) when compared with cases with only posterior teeth affected. In addition, the CT genotype of the tuftelin rs3790506 marker was overrepresented in cases with dmft scores higher than 5 \((p = 0.05)\) and dmfs scores higher than 6 \((p = 0.05)\) compared to controls, in cases with dmfs scores between 6 and 10 \((p = 0.03)\) compared to controls, and in cases with dmft scores between 5 and 8 \((p = 0.05)\) and dmfs scores between 6 and 10 \((p = 0.03)\) compared to controls.

We used regression analysis to assess, among the 91 cases, gene-environment interactions and the effects of confounders such as age or gender. Step-wise approaches were used and S. mutans data were used as an interactive term (table 4 lists the models tested). This analysis demonstrated that age was not correlated with dmft or dmfs scores, but gender was \((\text{dmft: } p = 0.006, R^2 = 0.08; \text{dmfs: } p = 0.10, R^2 = 0.03)\) with females having more caries lesions than males. The region of the mouth contributed most to dmft/dmfs scores with both anterior and posterior teeth affected contributing to dmft \((p < 0.001, R^2 = 0.30)\) and dmfs \((p = 0.002, R^2 = 0.10)\). When modeling S. mutans, only the enamelin rs3796704 marker \((T\text{ allele, common allele})\) contributed to S. mutans \((p = 0.03): -1.77 + 1.55 \cdot \text{rs3796704.}\)

The best-fitting model incorporating interactions \((\text{dmfs/dmft} = \alpha + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \ldots + \beta_n \cdot X_n + \beta_1 \cdot X_1 \cdot X_2 + \ldots + e)\) were:

1. \[\text{dmfs} = 9.90 + 1.88 \cdot \text{mouth region} - 0.19 \cdot \text{gender} + 5.11 \cdot \text{mouth region} \cdot \text{gender}\]
2. \[\text{dmfs} = 10.88 + 1.75 \cdot \text{mouth region} - 0.18 \cdot \text{gender} - 0.59 \cdot \text{rs17878486} + 4.99 \cdot \text{mouth region} \cdot \text{gender}\]
3. \[\text{dmfs} = 8.83 + 1.63 \cdot \text{mouth region} - 0.35 \cdot \text{gender} + 0.84 \cdot \text{rs3790506} + 5.35 \cdot \text{mouth region} \cdot \text{gender}\]

1 In the presence of Streptococcus mutans.
2 Only in less severely affected cases.

#### Discussion

Our report further supports the notion that genetic variation in enamel formation genes contributes to caries susceptibility. Previous work [Slayton et al., 2005] showed that an association between genetic variation of tuftelin and caries could only be detected when the interaction with S. mutans levels was included in the model. We did find evidence for an association between the tuftelin rs3790506 marker and caries experience, but these results were not improved by including S. mutans in the model. However, our S. mutans assessment was based on the presence or absence of the bacterial DNA in our assay and not semiquantitative such as the assessment done by Slayton et al. [2005]. A possible interaction between genetic variation and S. mutans could be seen for the enamelin marker \((p = 0.03)\).

Our study has some obvious limitations. We did not have bitewing radiographs, and incipient interproximal lesions may have not been detected in both the case and control groups. Cases and controls were not perfectly matched. There were slightly more females than males among the cases (although this difference was not statistically significant), and controls were on average 1 year older than the cases. Caries experience (dmft and dmfs scores) tends to increase with age; however, we do not think the difference in the ages between cases and controls greatly impacted our results since controls were older than cases. Finally, population substructure may have gone undetected. In Istanbul, 75% of the population is Turkish, and the remainder is represented by Kurds, Armenians, Jews and Greeks.

Our previous work with children older than 12 years and adults from Guatemala suggested that variation in amelogenin contributed to high caries susceptibility [Deeley et al., 2008]. It is remarkable to see that we could replicate this result in a sample population with several
differences: (1) population of Turkey versus Guatemala; (2) age groups (3- to 5-year-old children vs. individuals 12 or older), and caries measurement (dmft and dmfs including white-spot lesions vs. traditional DMFT). Both studies (in Guatemala and Turkey) had females more severely affected by caries than males. It is unanimous in the literature that females are more affected by caries than males [Lukacs and Largaespada, 2006]. Multivariable logistic regression analysis of data from more than 2,500 preschool children from Belgium [Declerck et al., 2008] showed that gender is significantly associated with caries experience, independently from age, home situation (having both parents at home vs. separated parents), educational level of parents, presence of dental plaque, age at start brushing, help with brushing, frequency of brushing, use of nursing bottle, application of sweet on pacifier, cleaning of pacifier in the parent’s mouth, snacking between meals and/or at night, and drinking sugared liquids between meals and/or at night. In addition, a genome-wide search of caries genetic susceptibility or protective loci in 46 Filipino families (in which females were more severely affected than males) also suggested a locus on chromosome Xq27.1 as harboring a protective caries gene (although this region is far from amelogenin) [Vieira et al., 2008]. The explanation for the gender bias seen in caries may lie in a genetic factor (or multiple genetic factors) on chromosome X.

Amelogenin is a likely candidate gene for caries susceptibility in humans and has differential expression in males versus females. Mutations and deletions in amelogenin cause one of the forms of X-linked amelogenesis imperfecta [reviewed by Hu et al., 2007]. However, amelogenesis imperfecta can be expressed differently depending on the gender of the affected individual. The enamel of affected males can be extremely thin and smooth, whereas in the teeth of affected females, enamel can be of almost normal thickness with defective vertical ridging [reviewed by Alvesalo, 1997]. Both the X and Y chromosomes have a version of the amelogenin gene [Nakahori et al., 1991]. However, the amino acid sequences of both X and Y amelogenin genes seem to differ to some extent, and the transcription products of the X and Y chromosomes are both quantitatively and qualitatively different [reviewed by Alvesalo, 1997]. The Y chromosome locus encodes a functional protein even though its level of expression is only 10% of that of the locus on the X chromosome [Salido et al., 1992]. One can speculate that variation of amelogenin X alters enamel structure and increases susceptibility to caries and that the additional 10% amelogenin expression in males partly explains why females tend to show higher scores of caries experience.

Our study also suggested for the first time that ameloblastin may contribute to caries susceptibility. Ameloblastin is expressed during the differentiation of inner enamel epithelium into ameloblasts, with intense localization in the Tomes processes of secretory ameloblasts. In contrast to amelogenin, only modest amounts of ameloblastin can be detected in enamel matrix [Krebsbach et al., 1996].

The cumulative evidence from our study and from the studies of others [Slayton et al., 2005; Deeley et al., 2008] strongly suggests a role of the genes involved in the enamel formation in caries susceptibility in humans. We hypothesize that genetic variation in these genes contributes to structural alterations of the enamel that may cause higher levels of mineral losses under acidic conditions and/or facilitate bacterial attachment and biofilm deposition. However, there is some variation among the studies regarding the genetic associations demonstrated (table 5). These differences can be the result of numerous factors such as heterozygosity of the genetic markers in the studied populations and variations in the study designs (data on primary vs. permanent dentitions, caries experience scores by tooth vs. by surface, availability of S. mutans data). Future studies will focus on the identification of functional variants within these genes that may provide a clear biological rationale on how variation in genes controlling enamel formation can influence future susceptibility to caries.

In summary, our results suggest that variation in amelogenin, ameloblastin and tuftelin contribute to caries susceptibility. In addition, variation in enamelin may interact with the presence of S. mutans infection.

Acknowledgements

The authors would like to thank the individuals that participated in this study for their support. This research is supported by NIH/NIDCR grant R01-DE014899 (to M.L.M.).
References
