



# Evolution of Cariogenic Character in *Streptococcus mutans*: Horizontal Transmission of Glycosyl Hydrolase Family 70 Genes

SUBJECT AREAS:  
ENVIRONMENTAL  
MICROBIOLOGY  
MICROBIAL GENETICS  
PATHOGENS  
MOLECULAR EVOLUTION

Tomonori Hoshino<sup>1</sup>, Taku Fujiwara<sup>1</sup> & Shigetada Kawabata<sup>2</sup>

Received  
4 May 2012

Accepted  
3 July 2012

Published  
18 July 2012

Correspondence and  
requests for materials  
should be addressed to  
T.F. (takufuji@  
nagasaki-u.ac.jp)

<sup>1</sup>Department of Pediatric Dentistry, Nagasaki University Graduate School of Biomedical Sciences, 1-7-1 Sakamoto, Nagasaki 852-8588, Japan, <sup>2</sup>Department of Oral and Molecular Microbiology, Osaka University Graduate School of Dentistry, 1-8 Yamadaoka, Suita, Osaka 565-0871, Japan.

Acquisition of the ability to produce polysaccharides from sucrose, i.e. the *gtf* gene encoding glucosyltransferase (GTF), is the key evolutionary event enabling dental biofilm formation by streptococci. To clarify the ancestry of streptococcal GTFs, time of its occurrence, and order of specific events, we investigated the distribution of GTFs among bacteria by phylogenetic analysis of the glycosyl hydrolase family 70 enzymes. We found that streptococcal GTFs were derived from other lactic acid bacteria such as *Lactobacillus* and *Leuconostoc*, and propose the following evolutionary model: horizontal gene transfer via transposons occurred when streptococci encountered lactic acid bacteria contained in fermented food. Intra-genomic gene duplication occurred by a secondary selection pressure such as consumption of refined sugar. Our findings concerning this evolution in *Streptococcus mutans* provide an important background for studies of the relationship between the historical spread of dental caries and anthropological factors.

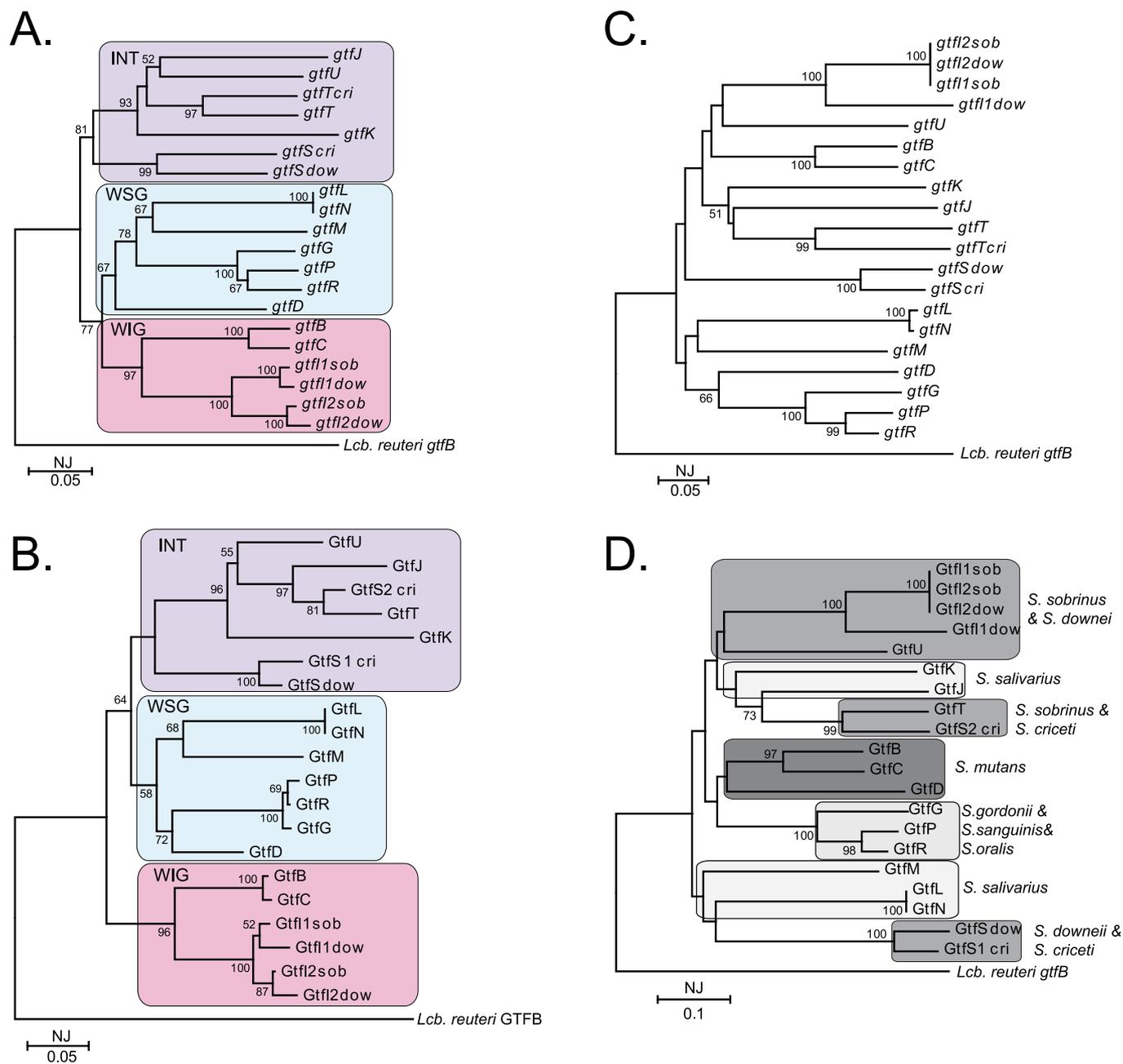
In humans, dental caries is defined as the demineralisation of teeth by the bacterial production of organic acid from fermentable substrates such as monosaccharides, oligosaccharides, and processed polysaccharides<sup>1-3</sup>. The dental caries pandemic has been previously discussed from an anthropological and archaeological viewpoints in terms of the consumption of fermentable carbohydrates<sup>4-7</sup>. Indeed, the emergence of dental caries positively correlates with the increased consumption of carbohydrates through the main source, grain<sup>8,9</sup>, following the advent of prehistoric farming. At that time, the degree of caries was comparatively mild, since oral bacteria produced only low amounts of organic acid. However, the widespread use of refined sugar during the industrial revolution due to improvements in food processing techniques, stabilisation of plantation-based sugar cane supplies, and the abolition of the Corn Law<sup>10</sup> led to a sudden increase in the prevalence of caries.

Thus far, only anthropological factors, like those described above, have been considered in the development of dental caries in humans, while factors affecting the disease-causing bacteria themselves have been largely ignored. This may be based on the firmly held belief that caries-causing bacteria preceded the development of caries. However, presently, there are no studies describing the origin of caries-causing bacteria. Today, *Streptococcus mutans* is considered the principal causative organisms of dental caries. These bacteria have the ability not only to metabolise large amounts of acid from various sugars but also to firmly adhere to the tooth surface using adhesive glucans produced from refined sugar by glucosyltransferases (GTFs)<sup>11</sup>. This latter property of caries-causing bacteria is the principal characteristic distinguishing prehistoric caries from modern caries.

Because of their role in the formation of dental biofilms, streptococcal GTF enzymes are thought to be one of the primary virulence factors of *Str. mutans* responsible for the development of dental caries<sup>1,11</sup>. Thus, phylogenetic analysis of the origin of GTFs may provide novel insights into the acquisition of cariogenicity by *Str. mutans*. GTFs (EC: 2.4.1.5), which are encoded by *gtf* genes, belong to the glycosyl hydrolase family 70 and catalyse the transfer of D-glucopyranosyl units from sucrose to acceptor molecules. To predict the ancestry of streptococcal GTFs, phylogenetic trees for the putative amino acid sequences of glycosyl hydrolase family 70 proteins were constructed using the neighbour-joining (NJ), minimum evolution (ME), and maximum parsimony (MP) methods. The surrounding sequences of the genes encoding the glycosyl hydrolase family 70 proteins were also analyzed. Our findings suggest that the genus *Streptococcus* acquired the *gtf* genes via







**Figure 3 | Phylogenetic analysis of the functional regions of streptococcal GTF.** Phylogenetic trees based on the gene encoding the catalytic region (A), the deduced amino acid sequence of the catalytic region (B), the gene encoding the first unit of the glucan-binding region constructed of 6 direct repeat units (C), and the deduced amino acid sequence of the first unit of the glucan-binding region were constructed by the neighbor-joining (NJ) method. In A and B, the streptococcal *gtf* gene family can be classified into 3 clusters: the water-soluble glucan synthesizing group (WSG), the water-insoluble glucan synthesizing group (WIG), and the intermediate group (INT). In D, GTFs from same or closely related species could be divided into the same cluster. The value on each branch is the estimated confidence limit (expressed as a percentage) for the position of the branches, as determined by bootstrap analysis. Only values exceeding 50% are shown. The scale bar ([NJ] distance) represents a 5% difference in nucleotide sequence.

**Phylogenetic analysis of the catalytic domain in streptococcal GTFs.** To infer age associated with ecological events, the mean interpopulational evolutionary diversity in the sequences encoding the catalytic domain of glycosyl hydrolase family 70 genes among *Streptococcus*, *Leuconostoc* and *Lactobacillus* was estimated as 0.043% (S.E.; 0.003). Further, to identify the most evolved enzyme among the streptococcal GTFs, we reconstructed the phylogenetic tree with only the streptococcal GTF catalytic domains and their coding sequences using the NJ method. The catalytic domain of *Lcb. reuteri* GTFB and its coding sequences were used as a root (Fig. 3A and B). Trees based on both the gene and deduced amino acid sequences indicated that the

streptococcal GTF family could be classified into 3 clusters: the water-soluble glucan-synthesising group (WSG), the water-insoluble glucan synthesizing group (WIG), and the intermediate group (INT). The phylogenetic distances based on the catalytic domain DNA sequences are shown in Table 1. The distances between the root and each streptococcal *gtf* gene group were found to be almost identical. In addition, the distance between the root and *gtfB* was found to be the same as the distance between the root and *gtfD*. In contrast, the distance between the root and *gtfC* was slightly larger in comparison.

Further analysis of the streptococcal GTF catalytic domains using a BLAST search showed that these regions are homologous to



**Table 1 | Mean phylogenetic distances (lower triangle) and standard error (upper triangle) of the catalytic regions among *Lcb. reuteri* GTFB (LR GTFB), streptococcal GTF groups, and 3 *gtf* genes from *Str. mutans***

	root	WSG	WIG	INT	<i>gtfB</i>	<i>gtfC</i>	<i>gtfD</i>
LR GTFB		0.020	0.021	0.019	0.024	0.024	0.024
WSG	0.558		0.015	0.014			
WIG	0.555	0.366		0.015			
INT	0.571	0.393	0.381				
<i>gtfB</i>	0.539					0.017	0.021
<i>gtfC</i>	0.560				0.144		0.022
<i>gtfD</i>	0.537				0.397	0.338	

Abbreviations: WSG, water-soluble glucan-synthesizing group; WIG, water-insoluble glucan synthesizing group; INT, intermediate group (INT).

$\alpha$ -amylase, but not to sucrase. In the blastp analysis of the catalytic domain of *Str. mutans* GTFB, the E-value for the  $\alpha$ -amylase catalytic domain was 8.25e-03. This finding suggests that the catalytic regions of streptococcal GTFs are not phylogenetically related to sucrase, despite having the same function.

**Phylogenetic analysis of the glucan-binding domain in streptococcal GTFs.** Since the glucan-binding domains in streptococcal GTFs are closely associated with cariogenicity, phylogenetic analyses of their coding genes and putative amino acid sequences of the first unit of the glucan-binding domains constructed of 6 direct repeating units were also performed. The phylogenetic trees were constructed using the NJ method, and the glucan-binding domains of *L. reuteri* GTFB were used as a root (Fig. 3C and D). The trees based on both the coding genes and the putative amino acid sequences could not be classified into 3 clusters as shown in Figure 3A and B. Although the periodicity of clustering was not observed in the tree based on the genes, the tree based on the putative amino acid sequence indicated that GTFs from the same or closely related species could be divided into the same cluster. In the SYSTERS Protein Family Database, the glucan-binding regions were categorized as Pfam CW\_binding 1 (PF01473), which was a cell wall binding protein family that included glucan-binding protein (Gbp), choline-binding protein, 1,4- $\beta$ -N-acetylmuramidase, and N-acetylmuramoyl-L-alanine amidase ([http://systems.molgen.mpg.de/cgi-bin/nph-fetchcluster.pl?PFAM=CW\\_binding\\_1](http://systems.molgen.mpg.de/cgi-bin/nph-fetchcluster.pl?PFAM=CW_binding_1)).

## Discussion

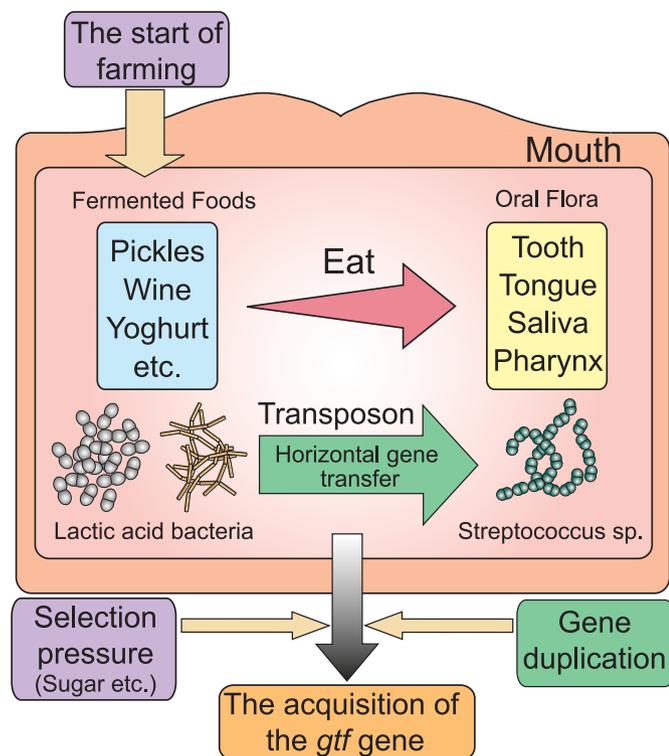
Phylogenetic analyses of glycosyl hydrolase family 70 enzymes were performed. Glucosyltransferase-S enzymes from *Lactococcus lactis*, encoded by *pspA* or *pspB*<sup>12</sup>, have the simplest structures and were consequently defined as the roots of the dendrograms. However, these enzymes do not actually belong to the glycosyl hydrolase family 70 enzymes, since they do not possess the catalytic and glucan-binding domains characteristic of this enzyme family shown in the alignment analysis. The constructed trees indicated that the streptococcal GTFs were derived from other lactic acid bacteria following their spread through the genera in the order of *Lactococcus*, *Lactobacillus*, *Leuconostoc*, and *Streptococcus*. More specifically, our data suggests that *Lcb. reuteri* GTFB or *Lcb. reuteri* GTFML4 would be the practical ancestor of the streptococcal GTFs. Our previous finding that only 51% of 41 *Str. oralis* clinical isolates possessed the *gtf* gene<sup>13</sup> supports the idea that some of the *Streptococcus* spp. are in the process of acquired the *gtf* gene. The average rate of sequence divergence at synonymous sites as determined by a comparison of homologous protein-coding regions, is 0.90% per million year<sup>14</sup>. Given that the interpopulational divergence in the sequences encoding the catalytic domain of GTFs is 0.043%, their expansion would have occurred approximately 48,000 years ago. Thus, *Streptococcus* acquired the gene only recently in bacterial evolutionary time.

It may seem likely that glycosyl hydrolase family 70 enzymes were produced by a combination of sucrase and glucan-binding proteins, such as Gbp from *Str. mutans*. However, *in silico* analyses using the DNA and protein databases, revealed that streptococcal GTF catalytic regions are homologous to  $\alpha$ -amylase, but not to sucrase. These analyses further showed that the glucan-binding regions are categorized as CW\_binding 1, including 1,4- $\beta$ -N-acetylmuramidase and N-acetylmuramoyl-L-alanine amidase. These findings suggest that the common ancestor of the glycosyl hydrolase family 70 enzymes would be an enzyme group whose general function was sugar hydrolysis and sugar chain binding. In addition, glucan-binding proteins, such as Gbp, were found in *Str. mutans* only in the present database. For this reason, it has been suggested that Gbp, one of the virulence factors allowing adherence to the tooth surface, was first produced due to the lack of the first half of glucosyltransferase in *Str. mutans*.

Homologues can be acquired through gene duplication (paralogs) or horizontal gene transfer (orthologs)<sup>15,16</sup>, and evidence exists to support both methods of acquisition in the glycosyl hydrolase family 70 genes. Our study located a transposase encoding gene upstream or downstream of the *gtf* gene in both *Lcb. reuteri* and *Lcb. sakei*<sup>17</sup> (Fig. 2A). A sequence highly homologous to streptococcal transposase genes was also found downstream of *Str. oralis gtfR*<sup>18</sup> (Fig. 2A and B). It has been reported that intra- and interspecies genomic variation in *Streptococcus* spp. is associated with transposase and that the bacterial population is a flexible gene pool<sup>19</sup>. These findings suggest that the streptococcal *gtf* gene was acquired through horizontal gene transfer via transposons, while periodic extinguishing of transposable elements may explain their absence around the *gtf* gene in some species of bacteria<sup>20</sup>. On the other hand, glycosyl hydrolase family 70 genes are believed to easily duplicate, since many bacteria possess multiple glycosyl hydrolase family 70 genes within the same genome. Tandem-arranged genes such as *gtfB* and *gtfC* (Fig. 2) are considered to have replicated by tandem duplication, a phenomenon observed in *Str. mutans*<sup>21</sup>, *Streptococcus salivarius*<sup>22</sup>, and *Streptococcus criceti* and in the lactic acid bacterium, *Lcb. reuteri* (Fig. 1). Based on these findings, *Str. mutans* may have acquired the 3 *gtf* genes as orthologs from other bacteria and/or as paralogs by duplication.

The acquisition of a gene is caused as the result of adaptation to habitat<sup>23</sup>. During this process, it is thought that some parts of the acquired gene are modified to adapt to the species-specific circumstances and that the other parts are conserved to keep the primary function<sup>14</sup>. For this reason, it is thought that the diversity of the conserved region reflects the systematic evolution of the acquired gene without species-specific modification. The catalytic regions of GTFs are highly homologous among streptococcal species and are highly conserved to keep sucrase activity. Thus, to observe the systematic evolution of streptococcal *gtf* genes, the phylogenetic analyses of catalytic region were carried out (Fig. 3A and B). Our results show that each phylogenetic distance between the root and the 3 streptococcal *gtf* gene groups (WIG, WSG, and INT) was almost identical, suggesting that these groups differentiated within the same period. This finding is in contrast with the commonly held view by dental researchers that GTFB and GTFC are the most advanced *S. mutans* GTF enzymes, since they primarily synthesise the water-insoluble  $\alpha$ -1,3-linked glucans directly associated with cariogenicity.

Here, we have shown that enzymes in the upper part of the phylogenetic tree synthesise glucans with various linkage types such as  $\alpha$ -1,3;  $\alpha$ -1,6;  $\alpha$ -1,2; and  $\alpha$ -1,4, while those in the lower part of the tree synthesise only water-soluble  $\alpha$ -1,6-linked glucans (Fig. 1 and Supplementary Table S1). Previous studies have shown that streptococcal GTF produces soluble glucans as a result of the displacement of amino acids in the catalytic domain and/or a decrease in the number of repeating units in the glucan-binding domain<sup>24,25</sup>. Thus, it was suggested that streptococcal GTF evolved to synthesise water-soluble glucans and that the ability to synthesise water-insoluble



**Figure 4 | Hypothetical model of GTF enzyme acquisition by oral *Streptococci*.** The genus *Streptococcus* acquired the *gtf* gene via horizontal gene transfer when it encountered lactic acid bacteria, such as *Lactobacillus* and *Leuconostoc*, in the human oral cavity. This event coincided with the introduction of fermented foods into the human diet, which was brought about by the development of agricultural practices and food preservation methods.

glucans would only be acquired if the ancestral character was strongly retained and highly specialised.

As shown in Figure 3A and B and Table 1, both *gtfB* and *gtfC* belong to the WIG cluster and adjoin each other in the phylogenetic tree, with only a short distance of 0.144 separating them. *gtfC* is probably replicated after *gtfB*, as the phylogenetic distances between the 2 genes and the root are 0.560 and 0.539, respectively. In contrast, *gtfB* and *gtfD* are divided into separate clusters, and the phylogenetic distance between them is relatively high (0.397). Thus, it is likely that these 2 genes were acquired separately; however, the similar phylogenetic distances between these genes and the root (*gtfB*, 0.539; *gtfD*, 0.537) suggest that they were acquired at similar time points.

Character acquisition is an evolutionary response that follows exposure to different selection pressures such as starvation conditions and thermal stress<sup>26</sup>. The resultant functional divergence plays an important role in the varying environmental adaptations of organisms<sup>23,27–29</sup>. For oral streptococci, it is conceivable that the availability of fermented foodstuffs and encounters with other lactic acid bacteria that possessed glycosyl hydrolase family 70 enzymes were major environmental selection pressures contributing to their acquisition of *gtf* genes. Non-oral streptococci, such as *Str. pneumoniae* and pyogenic group streptococci do not possess the *gtf* gene. At least, the *gtf* gene does not exist in the genome of those streptococci on the current database. This suggested that those streptococci that did not encounter the lactic acid bacteria carrying the *gtf* gene could not acquire the gene. We propose that *Streptococcus* obtained the first *gtf* gene via gene transfer as an initial step towards cariogenicity. The character acquisition was then completed following the consumption of refined sugar by humans, which acted as a secondary selection pressure prompting the acquisition of multiple *gtf* genes (Fig. 4). We also propose that this acquisition allowed *Str. mutans* to develop a

niche on the tooth surface, as opposed to other oral regions, such as the saliva, tongue, pharynx, or larynx.

In this study, we suggest that the reasons underlying the emergence of dental caries were not only anthropological but also bacteriological, i.e., genetic events leading to the evolution of cariogenic *Str. mutans*. We therefore recommend that a molecular biological approach, such as PCR-based detection of the *gtf* genes<sup>30</sup>, be adopted for future anthropological and archaeological studies. This would enable a more precise distinction between caries and tooth wear as well as analysis of the historical expansion of *Str. mutans*.

## Methods

**Glycosyl hydrolase family 70 sequences.** The DNA and amino acid sequences of glycosyl hydrolase family 70 proteins used in this study were obtained from GenBank at NCBI (<http://www.ncbi.nlm.nih.gov/>) with cross-reference to Pfam (<http://pfam.sanger.ac.uk/>). Sequencing analyses of unknown streptococcal *gtf* genes was performed as described previously<sup>18,30</sup>, and the sequences were deposited in the DNA Data Bank of Japan ([http://sakura.ddbj.nig.ac.jp/sakura\\_en.html](http://sakura.ddbj.nig.ac.jp/sakura_en.html)). We analysed 20 GTFs from *Streptococcus*; 2 GTFs, 9 dextran sucrases, and 1 alternan sucrose from *Leuconostoc*; 10 glucan sucrases from *Lactobacillus*; and 2 GTFs from *Lactococcus*. The NCBI accession numbers of these GTFs are provided in Figure 1 and Supplementary Table S1.

**Homology and protein family search.** Motifs from GTFs were analyzed by homology and protein family searches using BLAST in NCBI and SYSTERS Protein Family Database in Max Plank Institute for Molecular Genetics, Computational Molecular Biology (<http://systers.molgen.mpg.de/>), respectively.

**Phylogenetic analysis.** Sequence alignment was performed using ClustalX software version 1.83<sup>31</sup> (<http://bips.u-strasbg.fr/fr/Documentation/ClustalX/#G>). Multiple alignment files saved by ClustalX in the Clustal format (\*.aln) were converted to the MEGA format (\*.meg) using the MEGA version 5 software<sup>32</sup> (<http://www.megasoftware.net/>). Phylogenetic analysis was performed by the NJ, ME, and MP methods using MEGA version 5 software. Phylogenetic distances were calculated by the NJ method using the same software. The corresponding parameter of the NJ algorithm was set at 'complete deletion', and the 'nucleotide: p-distance' model and 'bootstrap method' were used. The mean interpopulational evolutionary diversity of the sequences encoding the catalytic domain of GTFs among *Streptococcus*, *Leuconostoc*, and *Lactobacillus* was also calculated with MEGA 5. The 'complete deletion', 'bootstrap method', and 'p-distance' model were used in this analysis.

**Mapping of the genomic location of glycosyl hydrolase family 70 genes.** The locations of the genes encoding glycosyl hydrolase family 70 proteins from *Str. mutans*, *Lcb. reuteri*, and *Lcb. sakei* were obtained using the Annotation Search Tool in the Comprehensive Microbial Resource of J. Craig Venter Institute (<http://www.jcvi.org/>) or by referring to previous reports<sup>17,33</sup>. The upstream and downstream sequences of the *gtfB* gene from *Str. oralis* Uo5 genome (GenBank: FR720602.1) were analysed using BLAST (NCBI; <http://www.ncbi.nlm.nih.gov/BLAST/>).

1. Caufield, P. W. Dental caries: an infectious and transmissible disease. *Compend. Contin. Educ. Dent.* **26**, 10–16 (2005).
2. Miller, W. D. The microorganisms of the human mouth. *Am. J. Med. Sci.* **101**, 159 (1981).
3. Touger-Decker, R. & van Loveren, C. Sugars and dental caries. *Am. J. Clin. Nutr.* **78**, 881S–892S (2003).
4. Turner, C. G. Dental anthropological indications of agriculture among the Jomon people of Central Japan. *Am. J. Phys. Anthropol.* **51**, 619–635 (1979).
5. Lukacs, J. R. Dental paleopathology and agricultural intensification in south Asia: new evidence from Bronze Age Harappa. *Am. J. Phys. Anthropol.* **87**, 133–150 (1992).
6. Nelson, G. C., Lukacs, J. R. & Yule, P. Dates, caries, and early tooth loss during the Iron Age of Oman. *Am. J. Phys. Anthropol.* **108**, 333–343 (1999).
7. Eshed, V., Gopher, A. & Hershkovitz, I. Tooth wear and dental pathology at the advent of agriculture: new evidence from the Levant. *Am. J. Phys. Anthropol.* **130**, 145–159 (2006).
8. Larsen, C. S. & Thomas, D. H. in *Anthropological Papers of the American Museum of Natural History Vol. 57 Anthropological papers of the American Museum of Natural History*. v. 57, pt. 3, 157–270 (American Museum of Natural History, New York, 1982).
9. Larsen, C. S. Behavioural implications of temporal change in cariogenesis. *Annu. Rev. Anthropol.* **24**, 185–213 (1983).
10. Yudkin, J. Sugar and disease. *Nature* **239**, 197–199 (1972).
11. Kuramitsu, H. K. Virulence factors of mutans streptococci: role of molecular genetics. *Crit. Rev. Oral. Biol. Med.* **4**, 159–176 (1993).
12. Bolotin, A. et al. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res.* **11**, 731–753 (2001).



13. Hoshino, T., Fujiwara, T. & Kilian, M. Use of phylogenetic and phenotypic analyses to identify nonhemolytic streptococci isolated from bacteremic patients. *J. Clin. Microbiol.* **43**, 6073–6085 (2005).
14. Ochman, H., Elwyn, S. & Moran, N. A. Calibrating bacterial evolution. *Proc. Natl. Acad. Sci. U S A* **96**, 12638–12643 (1999).
15. Ochman, H., Lawrence, J. G. & Groisman, E. A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304 (2000).
16. Jain, R., Rivera, M. C., Moore, J. E. & Lake, J. A. Horizontal gene transfer accelerates genome innovation and evolution. *Mol. Biol. Evol.* **20**, 1598–1602 (2003).
17. Kralj, S. *et al.* Glucan synthesis in the genus *Lactobacillus*: isolation and characterization of glucansucrase genes, enzymes and glucan products from six different strains. *Microbiology* **150**, 3681–3690 (2004).
18. Fujiwara, T., Hoshino, T., Ooshima, T., Sobue, S. & Hamada, S. Purification, characterization, and molecular analysis of the gene encoding glucosyltransferase from *Streptococcus oralis*. *Infect. Immun.* **68**, 2475–2483 (2000).
19. Hakenbeck, R. *et al.* Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. *Infect. Immun.* **69**, 2477–2486 (2001).
20. Wagner, A. Periodic extinctions of transposable elements in bacterial lineages: evidence from intragenomic variation in multiple genomes. *Mol. Biol. Evol.* **23**, 723–733 (2006).
21. Fujiwara, T. *et al.* Molecular analyses of glucosyltransferase genes among strains of *Streptococcus mutans*. *FEMS Microbiol. Lett.* **161**, 331–336 (1998).
22. Simpson, C. L., Cheetham, N. W., Giffard, P. M. & Jacques, N. A. Four glucosyltransferases, GtfJ, GtfK, GtfL and GtfM, from *Streptococcus salivarius* ATCC 25975. *Microbiology* **141** (Pt 6), 1451–1460 (1995).
23. Marri, P. R., Hao, W. & Golding, G. B. Gene gain and gene loss in streptococcus: is it driven by habitat? *Mol. Biol. Evol.* **23**, 2379–2391 (2006).
24. Monchois, V., Vignon, M., Escalier, P. C., Svensson, B. & Russell, R. R. Involvement of Gln937 of *Streptococcus downei* GTF-I glucansucrase in transition-state stabilization. *Eur. J. Biochem.* **267**, 4127–4136 (2000).
25. Abo, H. *et al.* Peptide sequences for sucrose splitting and glucan binding within *Streptococcus sobrinus* glucosyltransferase (water-insoluble glucan synthetase). *J. Bacteriol.* **173**, 989–996 (1991).
26. Tenaillon, O., Denamur, E. & Matic, I. Evolutionary significance of stress-induced mutagenesis in bacteria. *Trends Microbiol.* **12**, 264–270 (2004).
27. Hooper, S. D. & Berg, O. G. Duplication is more common among laterally transferred genes than among indigenous genes. *Genome Biol.* **4**, R48 (2003).
28. Hooper, S. D. & Berg, O. G. On the nature of gene innovation: duplication patterns in microbial genomes. *Mol. Biol. Evol.* **20**, 945–954 (2003).
29. Marri, P. R., Hao, W. & Golding, G. B. The role of laterally transferred genes in adaptive evolution. *BMC Evol. Biol.* **7** Suppl 1, S8 (2007).
30. Hoshino, T. *et al.* PCR detection and identification of oral streptococci in saliva samples using gtf genes. *Diagn. Microbiol. Infect. Dis.* **48**, 195–199 (2004).
31. Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* **23**, 403–405 (1998).
32. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739, doi:msr121 [pii] 10.1093/molbev/msr121 (2011).
33. Ajdić, D. *et al.* Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. *Proc. Natl. Acad. Sci. U S A* **99**, 14434–14439 (2002).

## Acknowledgments

The authors would like to thank Ikuri Konishi for providing technical assistance. This work was supported by KAKENHI (Grant-in-Aid for Scientific Research) from the Japan Society for the Promotion of Science (no. 23659968 and 24659912) and the JSPS International program for Young Researcher Oversea Visits.

## Author contributions

T.H. conceived the study. S.K. and T.F. supervised this project. T.H. analyzed the genetic data and constructed phylogenetic trees. S.K. and T.F. collected the genetic data and phylogenetic trees. T.H. and T.F. wrote the paper. All authors discussed the results and edited the manuscript.

## Additional information

**Accession code:** The datum of *Str. sanguinis gtfP* has been deposited in the GenBank database under accession code AB252650.

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**License:** This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

**How to cite this article:** Hoshino, T., Fujiwara, T. & Kawabata, S. Evolution of Cariogenic Character in *Streptococcus mutans*: Horizontal Transmission of Glycosyl Hydrolase Family 70 Genes. *Sci. Rep.* **2**, 518; DOI:10.1038/srep00518 (2012).