IJHSR
ijhighschoolresearch.org
■ RESEARCH ARTICLE

Gene Enrichment Analysis of Dental Cavities Associated Human Genes

Jeffrey Sun

Korea International School, Daewangpangyo-ro, 385 beon-gil, Bundang-gu, Seongnam-si, Gyeonggi-do, South Korea, gosyber1@gmail.com

ABSTRACT: Dental cavities cause major and prevalent public health problems. Previously, system-wide profiling of genes and proteins in human cells identified 47 human genes that are significantly associated with dental cavities. However, these genes need to be further analyzed for their collective functions to explore a new linkage on dental cavities. We performed gene enrichment analysis using the Enrich R program and analyzed gene network analysis using the GeneMania program in this study. Among 47 genes, we identified three genes (*WNT10A*, *KCNJ2*, and *FGF10*) associated with the abnormal development of primary teeth through gene enrichment analysis. Also, gene network analysis of *WNT10A*, *KCNJ2*, and *FGF10* showed that these genes are significantly associated with G protein-coupled receptor binding genes, regulating oral tissue and tooth development. Overall, elucidation of the gene network regulating the development of the dental cavity and abnormal development of primary teeth will provide insight into the etiology underlying these disorders. Also, this study provides potentially identifies novel therapeutic targets and contributes to regenerative medicine.

KEYWORDS: Biology; Genetics; Dental Cavity; Gene Enrichment analysis; Enrich R; GeneMania.

■ Introduction

Dental cavities, one of the most common health problems for children and adolescents, are caused by bacteria in our mouth called *Streptococcus mutans*.[1] The presence of these mutans results in biofilm formation, which lowers the pH, eventually creating cavities.[2] If cavities are not treated quickly and appropriately, they may enlarge in size, leading to more severe results. Symptoms caused by cavities vary from head/toothaches to visible holes forming.[3]

Dental cavities are more visible in younger children, as 52% of children aged six to eight had cavities. 25% of adults aged 20-64 only possess one cavity.[4] In addition, according to a study conducted by the Centers for Disease Control and Prevention in 2016, over 90% of adults in the United States have experienced a cavity.[5] Surprisingly, children of uneducated and low-income families are twice as likely to have cavities than children in high-income families.[6]

Genome-wide association studies (GWAS) for dental cavities have discovered the genetic risk factor, including overall cavities experienced and the presence or absence of disease in human populations.[7] However, only a few genetic-association markers for dental cavities and periodontitis have been found in children. Also, since the previous studies have not shown consistent evidence of specific genetic contributions to periodontitis and dental cavities, further research is needed to understand the potential cause of delta cavity development.[8]

System-wide profiling of genes and proteins in human cells produces a list of genes that are significantly associated with dental cavities. However, these genes need to be further analyzed for their collective functions to explore a new linkage on dental cavities. A previous paper identified 47 novel risk loci for dental caries indicating genetic risk factors of dental caries.[9] Therefore, in this research, the list of 47 genes was further analyzed using Enrichr to rank terms from gene-set libraries by comparing the Fisher exact test. Overall, we performed gene enrichment analysis of these 47 genes to improve our understanding of potential causes and consequences of dental cavity development.

■ Methods

Gene enrichment analysis using EnrichR:

Previously identified 47 novel risk loci provide the genetic risk factor of dental caries. We used this gene list and input as the query genes in the EnrichR program. Human phenotype analysis provides the top ten terms significantly associated with 47 genes. The term, number of overlapping genes, p-value, adjusted p-value, and gene information was analyzed through enrichment analysis using the EnrichR program.

Gene network analysis using GeneMania:

WNT10A, *KCNJ2*, and *FGF10* genes were further analyzed using GeneMania to predict the function of these genes. GeneMANIA extends the list with functionally similar genes that it identifies using available genomics and proteomics data. GeneMANIA analyzed gene networks with physical protein networks, co-expressed genes, co-localization genes, shared biological pathways and shared protein domains.

■ Results and Discussion

Table 1: Previously identified 47 novel risk alleles for dental caries indicating genetic risk factors of dental caries.

Gene	rsID number	Chromosome number: Genome Position
AMY1C	rs72694438	1:104364878
KRTCAP2	rs4971099	1:155155608
SYT14	rs2046850	1:210304319

DOI: 10.36838/v5i3.10

ITPKB	rs3820640	1:226868918
KCNJ3	rs2652452	2:155670203
ZNF804A	rs263771	2:185921692
WNT10A	rs121908120	2:219755011
ADCY3	rs11676272	2:25141538
ALK	rs80270335	2:29616655
FAM150B	rs62106258	2:417167
AAK1	rs5831974	2:69704336
STAG1	rs61790808	3:136443008
KCNH8	rs9831002	3:18852697
OPA1	rs185566659	3:193394725
RARB	rs7429279	3:25118637
RBM5	3:50135699DI	3:50135699
EFNA5	rs1352724	5:107083487
C5orf66	rs1122171	5:134509987
CDH9	rs55769264	5:26928047
FGF10	rs1482698	5:44539453
HLA	rs9366651	6:26336696
LOC157273	rs898797	8:9229689
PBX3	rs10987008	9:128661600
DMRTA1	rs10811723	9:22542285
PRUNE2	rs7852129	9:79346204
STFA1P	rs7918807	10:10020194
P2RY2	rs149467613	11:72943483
KLRAP1	rs10772314	12:10704350
CA12	rs72748935	15:63639416
NEO1	rs6495046	15:73353175
CHRNA3	rs10851907	15:78915864
RHCG	rs2072693	15:90014945
TMEM219	rs8054556	16:29958216
SALL1	rs1108343	16:51211595
FOXL1	rs10048146	16:86710660
NPEPPS	rs3865314	17:45669524
HOXB-AS2	rs9905793	17:46635649
KCNJ2	rs34559440	17:68399112
LOC100499467	rs7217268	17:70338127
BAHCC1	rs57067187	17:79361332
MC4R	rs28822480	18:57924823
CRLF1	rs2238651	19:18718846
MAMSTR	rs11672900	19:49220323
MTMR3	rs140357883	22:30292811
FAM118A	rs1569414	22:45727565
HDX	rs5922945	23:83523015

An allele is a variant form of a gene. Some genes have a variety of different forms located at the same location on the chromosome, that is, on the gene locus. Humans are called diploid organisms because there are two alleles in each gene locus, and one allele is inherited from one parent. Sometimes, genetic modification prevents one or more proteins from working properly. Mutations can cause a protein to malfunction or not be produced by altering the gene's instructions. Previous research identified 47 novel risk alleles for dental caries (**Table 1)**.[9] These risk alleles are in many different chromosomal positions: Chromosome 1 to chromosome 23. With this gene list, we further predict the function of these genes using gene enrichment analysis and gene network analysis.

Table 2: Gene enrichment analysis results on human phenotype ontology.

Rank	Term	Overlap	P-value	Adjusted P-value	Genes
1	Abnormality of primary teeth (HP:0006481)	3/33	6.31E-05	0.015	*WNT10A* *KCNJ2* *FGF10*
2	Partial duplication of thumb phalanx (HP:0009944)	2/8	1.49E-04	0.016	SALL1 *FGF10*
3	Partial duplication of the phalanx of hand (HP:0009999)	2/9	1.92E-04	0.016	SALL1 *FGF10*
4	Abnormality of incisor morphology (HP:0011063)	2/12	3.51E-04	0.021	*WNT10A* *FGF10*
5	Duplication of thumb phalanx (HP:0009942)	2/14	4.83E-04	0.023	SALL1 *FGF10*
6	Conical tooth (HP:0000698)	2/15	5.56E-04	0.023	*WNT10A* *FGF10*
7	Periauricular skin pits (HP:0100277)	2/19	9.01E-04	0.028	SALL1 *KCNJ2*
8	Preauricular pit (HP:0004467)	2/19	9.01E-04	0.028	SALL1 *KCNJ2*
9	Preaxial foot polydactyly (HP:0001841)	2/23	0.00132	0.032	SALL1 *FGF10*
10	Abnormality of the antihelix (HP:0009738)	2/24	0.00144	0.032	SALL1 *FGF10*

Gene Ontology has developed an extensive ontology describing molecular functions, biological processes, and human diseases.[10] We performed a gene enrichment analysis of 47 genes based on human phenotype ontology (HPO) to investigate the novel link with human disease. Each term in the HPO describes a clinical abnormality **(Table 2)**. The most significant HPO is the abnormality of primary teeth, with an adjusted p-value of 0.015. This result is consistent with the previous study that abnormal dental structures (chipped, cracked, and impacted teeth) enhance the dental cavity.[11]

Three genes (*WNT10A, KCNJ2*, and *FGF10*) are associated with the development of abnormalities in primary teeth. *WNT10A* is a protein-coding gene that encodes secreted signaling proteins. Recently, a novel mutation in WNT10A caused oligodontia, a rare genetic disorder representing the absence of more than six teeth in the primary tooth in children. [12] Protein modeling and functional analysis were devised in a zebrafish model to determine the impacts of oligodontia. Examining the embryos showed that *WNT10A* is visible in the craniofacial region during crucial periods of tooth development. This prohibited normal tooth development at five days post-fertil-

DOI: 10.36838/v5i3.10

ization.[13] In sum, a variant of *WNT10A* is the driving factor of human oligodontia and failure of tooth development in zebrafish.

Also, single nucleotide polymorphism (SNP) in the *KCNJ2* gene, which encodes the inward rectifier potassium channel two protein, has been implicated in the Pierre Robin sequence and Andersen-Tawil syndrome, which shows abnormalities in tooth development during infancy. This early abnormal tooth development, such as delayed formation and eruption of teeth, showed a higher risk of dental cavities than normal tooth development in infants.[14] Overall, *KCNJ2* is significantly associated with dental cavity development at an early age by causing abnormal primary tooth development.

FGF10 encodes fibroblast growth factor 10 protein. This protein is involved in mitogenic and cell survival activities and tooth development. The individuals with the T allele in FGF10 rs900379 were likelier to have larger teeth.[15] Also, this gene regulates the mouse, rat, and human crown formation in molars. *FGF10* disappears after the transitional stage from crown formation to root. In conclusion, *FGF10* is an essential gene controlling molars and crown formation development.

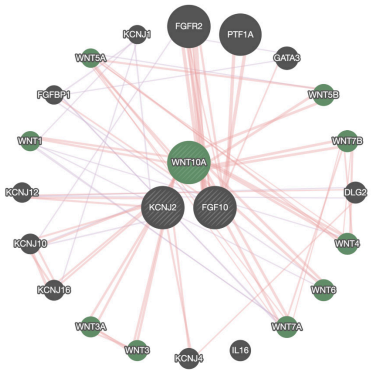

Figure 1: Gene network analysis of *WNT10A, KCNJ2,* and *FGF10* using GeneMANIA program. The green-colored genes are G protein-coupled receptor-binding genes. The red-colored line shows the physical interaction of protein. The purple line shows co-expressing genes.

GeneMANIA extends the list with functionally similar genes that it identifies using available genomics and proteomics data. GeneMANIA also weights the predictive value of each selected data set for the query. Therefore, we further analyzed the gene network of *WNT10A, KCNJ2*, and *FGF10* to predict the function of these genes using GeneMANIA. Figure 1 represents the gene network of *WNT10A*, *KCNJ2*, and *FGF10*, which is in the center. The functionally associated genes with *WNT10A*, *KCNJ2*, and *FGF10* were also represented around three genes **(Figure 1)**.

Interestingly, green-colored genes, including *WNT10A*, are functionally associated with G protein-coupled receptor binding genes. Previous research showed that G protein-coupled receptor Gpr115 is required for enamel mineralization mediated by ameloblasts. G protein-coupled receptors (GPCRs) function as transducers of external signals by activating associated G proteins and regulating cellular physiology.[16] Since tissue-specific GPCRs play essential roles in organ development, *WNT10A*-associated genes may regulate enamel mineralization meaning dysregulating *WNT10A* would be associated with the development of the dental cavity.

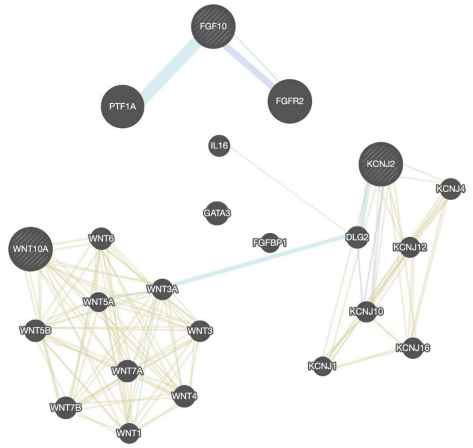

Figure 2: Gene network analysis of WNT10A, KCNJ2, and FGF10 indicates co-localization genes (sky blue), shared the same pathway (blue) and shared similar protein domains (yellow).

Since the protein sequences are often associated with gene function, we analyzed the genes that shared similar protein domains. The protein domain is an independent folding unit of the three-dimensional protein structure with functional importance. The yellow line indicates the connection between the genes that share the same domains **(Figure 2)**. *WNT10A* family genes contain a conserved sequence of the Wnt superfamily domain, which orchestrates the cellular processes, including cell proliferation, differentiation, and apoptosis.[17] *KCNJ2* family genes contain three domains: IRK_N superfamily, Ion_trans_2 superfamily, and IRK_C superfamily. This superfamily indicates that the *KCNJ2* family regulates the potassium ion channels in human cells.[18] *FGF10* is colocalized with PTF1A and shared pathway with FGFR2. *FGF10* gene encodes a protein fibroblast growth factor 10.[19] This protein may be colocalized with PTF1A and activate tissue repair by wound healing with the *FGFR2* gene. In sum, *WNT10A* and *KCNJ2* superfamily with the *FGF10* gene network is involved in the abnormal development of primary teeth leading to dental cavity formation.

Conclusion

There is a trade-off between phenotypic improvement and obtaining valid phenotypic data from a large population to allow for statistical power for genome-wide analysis. Given the logistical and economic difficulties of getting detailed phenotypes in large populations, this study has focused on genes associated with clinical outcomes indicative of the presence or absence of dental cavities.

Recent studies of other complex traits have taken a different approach, circumventing the limited availability of purified clinical phenotypes by utilizing very large population-based cohorts containing both genetic data and less purified proxy phenotypes. For example, the UK Biobank (UKB) contains information on approximately 500,000 participants. The existence of a less defined but informative phenotype has enabled the successful discovery of associative signals for several complex traits, providing critical biological insights. A deeper understanding of the genetic contribution to caries and

periodontitis requires a change in assay scale through similar measurement approaches.

Using previously discovered 47 novel risk loci for dental caries, we found that the heritability of dental caries is partially shared with a range of complex traits. Nine applied analyses using EnrichR and GeneMania suggested that *WNT10A, KCNJ2*, and *FGF10* are significantly associated with G protein-coupled receptor binding genes, regulating oral tissue and tooth development.

Characterizing the functional biology at these loci could provide insights that could improve diagnosis or guide the development of novel therapeutic approaches targeting the most prevalent global diseases. At the single variant and whole-genome level, the heritability of dental caries partially overlaps with other complex traits and diseases. Together, these findings improve our understanding of the potential causes and consequences of dental cavity development.

Acknowledgments

I want to give special thanks to Dr. Woo Rin Lee from the University of Suwon for his support.

References

1. Loesche, W. J. Role of Streptococcus Mutans in Human Dental Decay. *Microbiol. Rev.* **1986**, 50 (4), 353–380.
2. Matsui, R.; Cvitkovitch, D. Acid Tolerance Mechanisms Utilized by Streptococcus Mutans. *Future Microbiol*. **2010**, 5 (3), 403–417.
3. Peres, M. A.; Macpherson, L. M. D.; Weyant, R. J.; Daly, B.; Venturelli, R.; Mathur, M. R.; Listl, S.; Celeste, R. K.; Guarnizo-Herreño, C. C.; Kearns, C.; et al. Oral Diseases: A Global Public Health Challenge. *Lancet* **2019**, 394 (10194), 249–260.
4. Dye, B. A.; Hsu, K.-L. C.; Afful, J. Prevalence and Measurement of Dental Caries in Young Children. *Pediatr. Dent*. **2015**, 37 (3), 200–216.
5. Dye, B. A.; Mitnik, G. L.; Iafolla, T. J.; Vargas, C. M. Trends in Dental Caries in Children and Adolescents According to Poverty Status in the United States from 1999 through 2004 and from 2011 through 2014. *J. Am. Dent. Assoc.* **2017**, 148 (8), 550-565.e7.)
6. Lee, J. N.; Scott, J. M.; Chi, D. L. Oral Health Behaviours and Dental Caries in Low-Income Children with Special Healthcare Needs: A Prospective Observational Study. *Int. J. Paediatr. Dent.* **2020,** 30 (6), 749–757..
7. Witte, J. S. Genome-Wide Association Studies and Beyond. Annu. Rev. *Public Health* **2010**, *31*, 9-20 4 p following 20.
8. Alotaibi, R. N.; Howe, B. J.; Chernus, J. M.; Mukhopadhyay, N.; Sanchez, C.; Deleyiannis, F. W. B.; Neiswanger, K.; Padilla, C.;Poletta, F. A.; Orioli, I. M.; et al. Genome-Wide Association Study (GWAS) of Dental Caries in Diverse Populations. *BMC Oral Health* **2021**, *21* (1), 377.
9. Shungin, D.; Haworth, S.; Divaris, K.; Agler, C. S.; Kamatani, Y.; Keun Lee, M.; Grinde, K.; Hindy, G.; Alaraudanjoki, V.;Pesonen, P.; *et al*. Genome-Wide Analysis of Dental Caries and Periodontitis Combining Clinical and Self-Reported Data. *Nat. Commun*. **2019**, *10* (1), 2773.
10. The Gene Ontology Consortium. The Gene Ontology Resource: 20 Years and Still GOing Strong. *Nucleic Acids Res*. **2019**, *47* (D1) , D330–D338..
11. Costa, F. S.; Silveira, E. R.; Pinto, G. S.; Nascimento, G. G.; Thomson, W. M.; Demarco, F. F. Developmental Defects of Enamel and Dental Caries in the Primary Dentition: A Systematic Review and Meta-Analysis. *J. Dent.* **2017**, *60*, 1–7.
12. Park, H.; Song, J.-S.; Shin, T. J.; Hyun, H.-K.; Kim, Y.-J.; Kim, J.-W. WNT10A Mutations Causing Oligodontia. *Arch. Oral Biol.* **2019**, *103*, 8–11.
13. Yuan, Q.; Zhao, M.; Tandon, B.; Maili, L.; Liu, X.; Zhang, A.; Baugh, E. H.; Tran, T.; Silva, R. M.; Hecht, J. T.; et al. Role of WNT 10A in Failure of Tooth Development in Humans and Zebrafish. Mol. Genet. *Genomic Med.* **2017**, *5* (6), 730–741.
14. Pillas, D.; Hoggart, C. J.; Evans, D. M.; O'Reilly, P. F.; Sipilä, K.; Lähdesmäki, R.; Millwood, I. Y.; Kaakinen, M.; Netuveli, G.; Blane, D.; *et al*. Genome-Wide Association Study Reveals Multiple Loci Associated with Primary Tooth Development during Infancy. *PLoS Genet*. **2010**, *6* (2), e1000856.
15. Marañón-Vásquez, G. A.; Vieira, A. R.; Dos Santos, L. V.; Cunha, A. S.; Weiss, S. G.; Araujo, M. T. de S.; Bolognese, A. M.; Scariot, R.; Küchler, E. C.; Stuani, M. B. S. FGF10 and FGF13 Genetic Variation and Tooth-Size Discrepancies. *Angle Orthod.* **2021**, *91* (3), 356–362.
16. Chiba, Y.; Yoshizaki, K.; Saito, K.; Ikeuchi, T.; Iwamoto, T.; Rhodes, C.; Nakamura, T.; de Vega, S.; Morell, R. J.; Boger, E. T.; *et al*. G Protein-Coupled Receptor Gpr115 (Adgrf4) Is Required for Enamel Mineralization Mediated by Ameloblasts. *J. Biol. Chem.* **2020**, *295* (45), 15328–15341.
17. Doolan, B. J.; Onoufriadis, A.; Kantaputra, P.; McGrath, J. A. WNT10A, Dermatology and Dentistry. *Br. J. Dermatol*. **2021**, *185* (6), 1105–1111.
18. Kimura, H.; Zhou, J.; Kawamura, M.; Itoh, H.; Mizusawa, Y.; Ding, W.-G.; Wu, J.; Ohno, S.; Makiyama, T.; Miyamoto, A.; *et al*. Phenotype Variability in Patients Carrying KCNJ2 Mutations. Circ. Cardiovasc. *Genet*. **2012**, *5* (3), 344–353.
19. Bhushan, A.; Itoh, N.; Kato, S.; Thiery, J. P.; Czernichow, P.; Bellusci, S.; Scharfmann, R. Fgf10 Is Essential for Maintaining the Proliferative Capacity of Epithelial Progenitor Cells during Early Pancreatic Organogenesis. *Development* **2001**, *128* (24), 5109–5117.

Authors

Jeffrey is a junior at Korea International School. He is interested in studying biochemical engineering and wishes to major in dentistry. He hopes this research paper will be the start of studying more refined biological concepts in relation to dentistry.

DOI: 10.36838/v5i3.10