



Identification and Characterization of Key Genes Associated with Amelogenesis

Tahsinul Haque¹ Fatema Akhter² Nourelhoda Alim³ Abdullah Nabhan⁴ Fawzia Al kahtani⁵
Abdullah Mohammed Sambawa⁶

¹Preventive Dental Sciences Department, College of Dentistry, Dar Al Uloom University, Riyadh, Saudi Arabia

²Surgical and Diagnostic Sciences Department, College of Dentistry, Dar Al Uloom University, Riyadh, Saudi Arabia

³Surgical and Diagnostic Sciences Department, College of Dentistry, Dar Al Uloom University, Riyadh, Saudi Arabia

⁴Surgical and Diagnostic Sciences Department, Collage of Dentistry, Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia

⁵Dental Public Health, Private Sector, Riyadh, Saudi Arabia

⁶Primary Healthcare Center, Ministry of Health, Riyadh, Saudi Arabia

Address for correspondence Tahsinul Haque, PhD, Preventive Dental Sciences Department, College of Dentistry, Dar Al Uloom University, Riyadh, Saudi Arabia (e-mail: mdhaque@dau.edu.sa).

Eur J Dent 2025;19:116–123.

Abstract

Objectives The identification of key genes associated with amelogenesis would be helpful in finding solutions to genetic disorders in oral biology. The study aimed to use *in silico* analysis to identify the key genes involved in tooth development associated with preameloblasts (PABs) and secretory ameloblasts (SABs).

Material and Methods The data was subjected to quality analysis and uniform manifold approximation and projection analysis. To examine the distribution of the genes and identify important upregulated loci, a *p*-value histogram, a quantile plot, a mean difference and mean-variance plot, and a volcano plot were generated. Finally, protein-protein interaction and gene enrichment analyses were performed to determine the ontology, relevant biological processes, and molecular functions of selected genes.

Results A total of 157 genes were found to be significant in the PAB versus SAB comparison. HIST1H31 revealed strong interaction with HIST1H2BM, and EXO1, ASPM, SPC25, and TTK showed strong interactions with one other. The STRING database revealed that NCAPG, CENPU, NUSAP1, HIST1H2BM, and HIST1H31 are involved in biological processes. NCAPG, CENPU, SPC25, ETV5, TTK, ETV1, FAM9A, NUSAP1, HIST1H2BM, and HIST1H31 are involved in cellular components.

Conclusion The TTK, NUSAP1, CENPU, NCAPG, FAM9A, ASPM, SPC25, and HIST1H31 genes demonstrate functions in cell division. These genes might play a role in ameloblast development. These results will be useful in developing new methods to stimulate ameloblast development, which is essential for tooth regeneration and tissue engineering. However, more research is required to validate the functions of these genes and the genes with which they interact. A wide variety of genetic, epigenetic, and exogenous signaling factors regulate these genes and pathways throughout development and differentiation, cell fate, and behavior.

Keywords

- ▶ amelogenesis
- ▶ genes
- ▶ genetics
- ▶ preameloblasts
- ▶ secretory ameloblasts
- ▶ tooth development

Introduction

Amelogenesis is the process by which enamel is formed by epithelial ameloblasts facing the odontoblast layer. More advanced odontoblasts and stratum intermedium cells initiate ameloblast differentiation via molecular signals. Amelogenins, enamellins, ameloblastins, and tuftelins, which are released into the extracellular space by ameloblasts, are the major proteins of the enamel matrix. Enamel changes into highly mineralized tissue, and as a result, ameloblasts consume debris.¹ Apoptosis kills approximately half of the ameloblasts during amelogenesis, and the remaining half dies after the process has been completed. Amelogenesis is a process of gradual differentiation governed by a number of molecular and morphogenetic events and is distinguished by the formation of enamel and the differentiation of ameloblasts. Enamel is made by epithelial cells called ameloblasts that fuse with the odontoblast layer in a process called amelogenesis. The process of amelogenesis ends when the tooth erupts, and there is no secondary or regenerative enamel development afterward. Therefore, amelogenesis is distinguished by the transition from preameloblast (PAB) to secretory ameloblast (SAB). Following the disintegration of the basement membrane during the bell stage of tooth formation,² induced by signals from the underlying dental mesenchymal compartment, PABs differentiate into SABs, thereby initiating the differentiation process. These extracellular matrix proteins are necessary for the formation and mineralization of enamel,³ and disruption of ameloblast protein production results in deformed enamel.⁴ Ameloblasts differentiate into short mature ameloblasts (MABs) that regulate enamel mineralization after laying down a full-thickness enamel matrix. MABs die by apoptosis⁵ following tooth eruption. Because adults lack easily available ameloblast-lineage cells, enamel cannot be restored or regenerated. Amelogenesis imperfecta (AI) is a collection of inherited diseases that cause changes in the structure and appearance of dental enamel and are frequently associated with changes in other intraoral or extraoral tissues. AI is a group of genetically determined disorders that have a relatively uniform effect on the structure and clinical appearance of the enamel on all or nearly all teeth and that can be linked to morphologic or biochemical abnormalities elsewhere in the body.⁶ Amelogenin is expressed in humans and cattle from genes on the X and Y chromosomes. In human males, amelogenin (AMELX) transcripts are expressed from the X chromosomal copy of the gene, while 10% are expressed from the Y chromosomal copy (AMELY). The ENAM gene in this region has been identified as a potential gene affected in the form of AI.⁷

The ENAM gene contains 10 exons and 9 introns. The ameloblastin (AMBN) gene is strongly expressed by ameloblasts and weakly expressed by odontoblasts and preodontoblasts, with moderate expression also observed in Hertwig's epithelial root sheath and odontogenic benign tumors with malignant behavior such as ameloblastoma. The gene encoding human enamelysin (matrix metalloproteinase-20, or MMP-20) resides on chromosome 11. Many

studies have found that MMP-20 is only present in teeth. A genetic mutation of MMP-20 was discovered in two family members affected with autosomal-recessive pigmentary hypomaturity. This mutation removes the splice acceptor at the 3' end of intron 6 and results in the production of a hypomature enamel product. The most likely outcomes appear faulty splicing with both defects introducing an upstream translation termination codon.⁸ The KLK4 gene, found on chromosome 19, is a member of the human tissue kallikrein gene family. KLK4 is involved in fragmenting smaller pieces of the tumor necrosis factor receptor-associated protein amelogenin cleavage product. In the case of KLK4, one mutation leads to a shortened protein that has 152 but lacks the S207 site, which is required for the enzyme to function; because of this aberrant enzyme activity, the enamel crystallites expand to standard length but only partial thickness.⁹

Amelogenesis is one of the important mechanisms in tooth development. Unraveling the role of genes associated with the mechanisms will open new avenues for diagnostic techniques and cures related to tooth disorders. *In silico* studies in which gene identification and functional genomics studies are performed before designing follow-up laboratory research, are very useful. In the present study, various bioinformatics tools were used to identify and characterize amelogenesis-related genes to unravel their roles in tooth development. Moreover, using bioinformatics tools to study genes associated with dental biology and health is frequent in the dental sciences. Identification of relevant molecules and signaling networks with greater precision will shed light on the mechanisms that govern enamel formation and lay the groundwork for elucidating the mechanisms that regulate the differentiation of PABs into SABs to better understand enamel regeneration. This study has provided new information that will be helpful in finding pragmatic treatments for genetic disorders associated with amelogenesis.

Materials and Methods

Data Mining and Quality Control Analysis

The data was retrieved from the Gene Expression Omnibus (GEO) data set series GSE59214. After collecting expression profiling using array data, the GEO2R Web tool was used for further analysis to identify genes differing in expression across various experimental conditions. GEO2R is an interactive resource for contrasting two or more sample groups within a GEO series to locate genes that are expressed differentially in differing experimental settings.

The findings have been presented in the form of a gene table in order of importance, a series of graphical plots to aid in the visualization of differentially expressed genes, and a quality assessment of the data set, using the Bioconductor project's GEO query¹⁰ and limma R packages.¹¹ Limma (linear models for microarray analysis) is a popular statistical technique for identifying differentially expressed genes using the R programming language. Statistical *p*-values were subjected to multiple-testing corrections in a wide range of experimental designs and data sources to reduce

the likelihood of false-positive results. For comparison, the data was divided into two groups, namely, PAB and SAB, and each group contained three replications. To normalize data¹² and visualize the distribution of the values in the selected samples, an expression box plot and a density plot were generated.¹³

Uniform Manifold Approximation and Projection

Analysis

Next, a uniform manifold approximation and projection (UMAP) plot was generated. UMAP is a technique of dimension reduction that can be used to better view the connections that exist between individual samples.¹⁴

Value Histogram

To analyze the distribution of all genes according to their *p*-values, a *p*-value histogram plot¹⁵ was generated.

Mean Difference and Mean-Variance Plots

A mean difference plot was constructed to display the results of a single comparison (one group of PABs compared to another group of SABs). A mean-variance plot was generated to analyze the variation in the data.¹⁶

Volcano Plot

To identify the top-upregulated genes, a volcano plot¹⁷ analysis was employed. The top 20 upregulated genes were selected according to the lowest adjusted *p*-values. The most dependable genes exhibited the lowest *p*-values (►Table 1).

Protein-Protein Interaction and Gene Enrichment Analysis

An investigation of protein-protein interactions¹⁸ was performed on the selected genotypes. The g:Profiler web was used for gene enrichment analysis¹⁹ to analyze the ontology, relevant biological processes, and molecular functions of selected genes.

Results

Data Normalization Quality Control

Because the expression data was normalized, all selected samples had the same value distribution. The box plots in ►Fig. 1A show the normalization of two samples. The first sample represents PABs with three replications, and the second sample represents SABs with three replications. The plots show the data variation between the PABs and the SABs with their

Table 1 Top upregulated genes

ID	Adjusted <i>p</i> -value	<i>p</i> -Value	Log FC	Gene symbol	Gene title
8151341	0.00178	0.000000107	4.357	TRPA1	Transient receptor potential cation channel subfamily A member 1
8138289	0.00584	0.00000114	3.912	ETV1	ETS variant 1
8056572	0.00584	0.00000158	3.749	SPC25	SPC25, NDC80 kinetochore complex component
7940147	0.00584	0.00000138	3.551	FAM111B	Family with sequence similarity 111 member B
7923086	0.00584	0.000000825	3.218	ASPM	Abnormal spindle microtubule assembly
7967993	0.00584	0.00000105	2.564	FGF9	Fibroblast growth factor 9
8124531	0.00584	0.00000151	2.367	HIST1H3I	Histone cluster 1, H3i
8117594	0.01494	0.0000056	3.903	HIST1H2BM	Histone cluster 1, H2bm
7909946	0.01494	0.00000844	3.592	FAM177B	Family with sequence similarity 177 member B
8052355	0.01494	0.00000775	3.241	EFEMP1	EGF containing fibulin-like extracellular matrix protein 1
8171260	0.01494	0.00000886	3.189	FAM9A	Family with sequence similarity 9 member A
8094278	0.01494	0.00000679	2.939	NCAPG	Non-SMC condensin I complex subunit G
8103932	0.01494	0.00000783	2.862	CENPU	Centromere protein U
7926728	0.01494	0.00000823	2.285	MYO3A	Myosin IIIA
7910997	0.01494	0.00000791	2.149	EXO1	exonuclease 1
8117426	0.01494	0.0000085	1.94	HIST1H2BH/// HIST1H2B H	Histone cluster 1, H2bh///histone cluster 1, H2bh
8101788	0.01528	0.00000964	2.355	UNC5C	unc-5 netrin receptor C
7982889	0.01651	0.0000109	3.443	NUSAP1	Nucleolar and spindle associated protein 1
8120838	0.01659	0.0000118	2.563	TTK	TTK protein kinase

Abbreviations: EGF, epidermal growth factor; ETS, erythroblast transformation-specific; FC, fold change; SMC, structural maintenance of chromosomes.

Note: The top 20 upregulated genes, selected according to the lowest adjusted *p*-values (< 0.05).

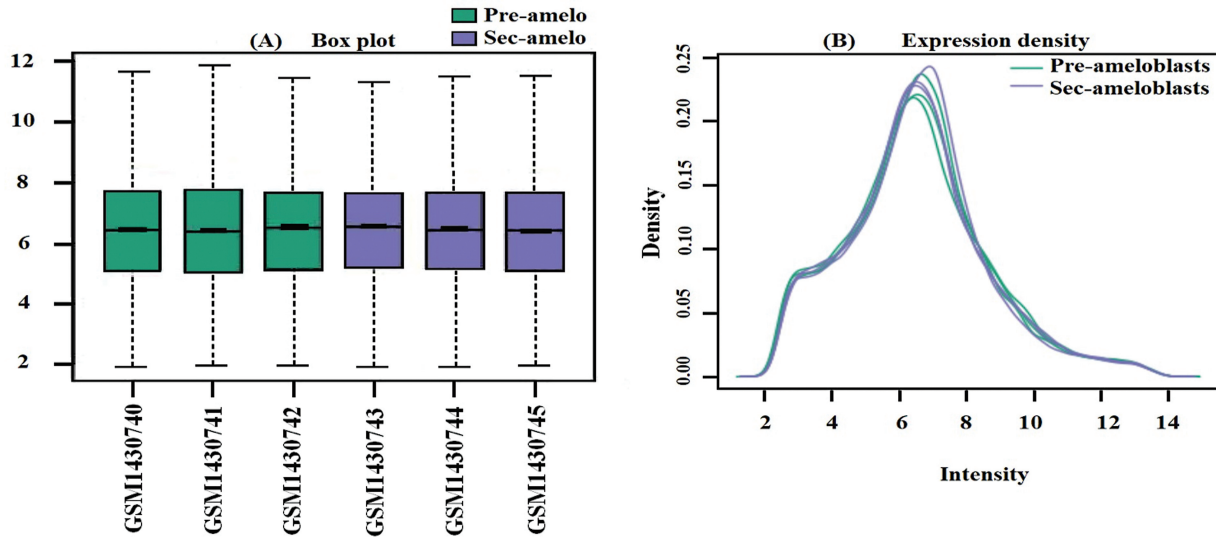


Fig. 1 (A) Diagnostic plots for microarray data from three samples of preameloblasts and three samples of secretory ameloblasts. (B) Density plot of preameloblast and secretory ameloblast samples.

replications. The density expression plot in ► **Fig. 1B** shows that the variation between the PABs and the SABs was small.

UMAP and p-Value Histogram Plot

The plot in ► **Fig. 2A** shows that there were substantial differences among thousands of genes in PABs and SABs. Variations among three biological replicates of PABs and three biological replicates of SABs were minimal; however, a distinct difference between samples of PABs and SABs was noted. The histogram in ► **Fig. 2B** displays the set of *p*-values. From the *p*-value histogram, it is evident that the *p*-values were not uniformly distributed.

Variance Trend and Mean Difference Plot

The graph shown in ► **Fig. 3A** illustrates the relationship between the mean and variance of the PAB and SAB expression

data and aided in determining whether there was a high level of variety in the data. This graph could also be used to determine whether the precision weights option for accounting for mean-variance trends should be used; precision weights improve test result precision when there is a strong mean-variance trend. Each dot corresponds to a gene. The mean difference plot of PABs versus SABs denotes upregulated genes by red dots and downregulated genes by blue dots, showing the average of normalized counts versus log2 fold changes for all genes studied.

Venn Diagram

A Venn diagram of the genes indicated that out of a total of 33,140 genes, 157 were significant in the PAB versus SAB comparison (► **Fig. 4**). Data regarding the significant genes has been provided in the Excels spreadsheet.

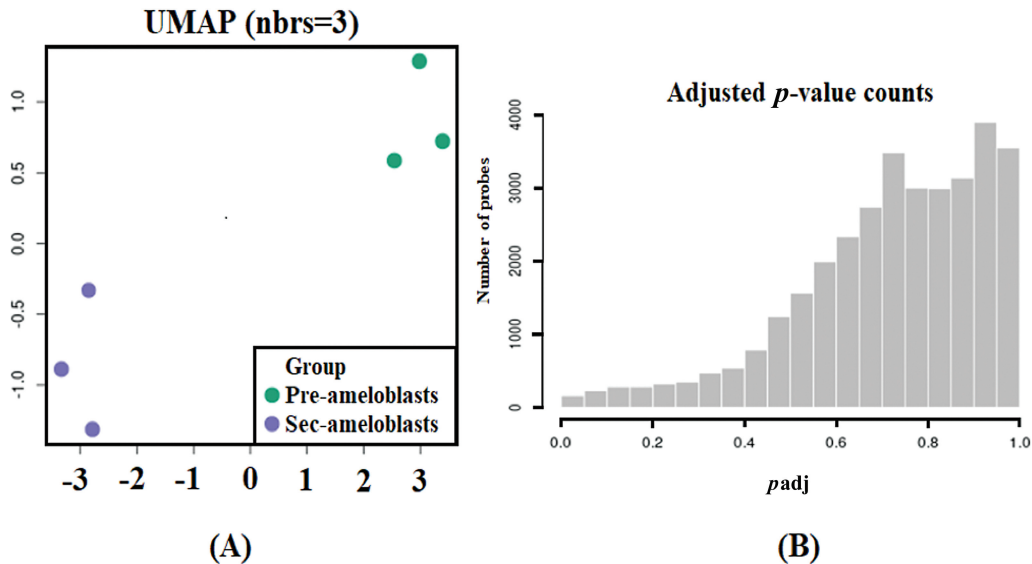


Fig. 2 (A) Uniform manifold approximation and projection (UMAP) plot generated from three samples of preameloblasts and three samples of secretory ameloblasts. (B) Adjusted *p*-value histogram showing the *p*-value distribution for all analyzed genes.

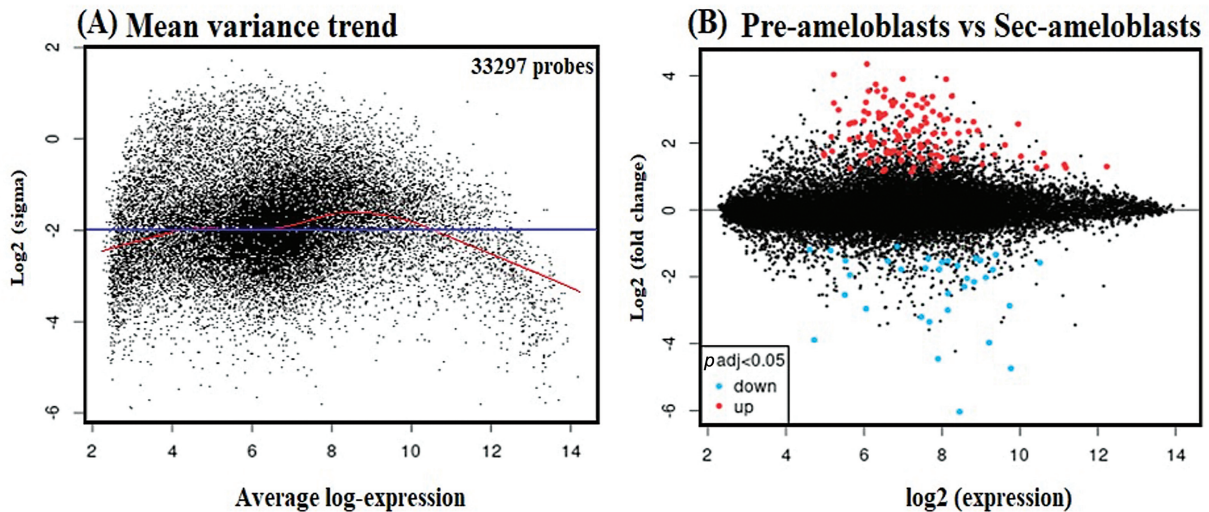


Fig. 3 (A) Variance trend plot of preameloblasts versus secretory ameloblasts. Each dot corresponds to a gene. The red line represents an approximation of the mean-variance trend. The blue line is an approximation to a constant variance. (B) Mean difference plot of preameloblasts versus secretory ameloblasts.

Volcano Plot

The volcano plot (►Fig. 5) of PABs versus SABs presents *p*-value versus magnitude of change statistical significance (fold change). It represents the identification of genes with large, statistically significant changes. In this volcano plot, the most upregulated genes are on the right, the most downregulated genes are on the left, and the most statistically significant genes are at the top.

The volcano plot was constructed to identify the top-upregulated genes in PABs versus SABs. The increase in upregulated genes was fourfold in the PAB versus SAB plot. All of the gene data is presented in an Excel spreadsheet.

Protein-Protein Interactions

Protein-protein interaction analysis was performed on the set of genes selected according to the adjusted *p*-values. According to this analysis, HIST1H31 revealed strong interaction with HIST1H2BM, and EXO1, ASPM, SPC25, and TTK

showed strong interaction with one other (►Fig. 6). The STRING database indicates that NCAPG, CENPU, NUSAP1, HIST1H2BM, and HIST1H3I are involved in biological processes. NCAPG, CENPU, SPC25, ETV5, TTK, ETV1, FAM9A, NUSAP1, HIST1H2BM, and HIST1H3I are related to cellular components (►Table 1).

Gene Enrichment Analysis

Functional enrichment analysis was performed on selected genes from PABs and SABs. This analysis provided information regarding the molecular function of gene ontology, biological process of gene ontology, and cellular components of gene ontology. The analysis revealed that the genes are involved in biological processes, cellular components, and KEGG pathways, as shown in ►Fig. 7. The table of gene ontology with *p*-value is shown in ►Fig. 8, and the functional descriptions are provided in the Excel data sheet.

GSE59214: limma, *p*_{adj}<0.05

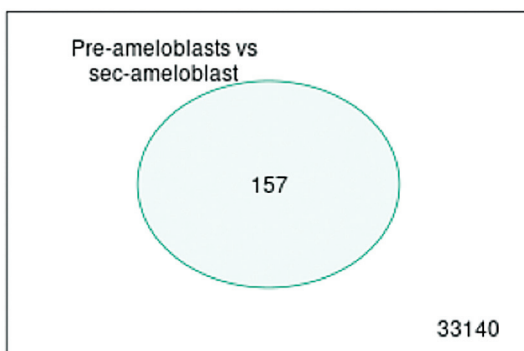


Fig. 4 Venn diagram of preameloblasts versus secretory ameloblasts.

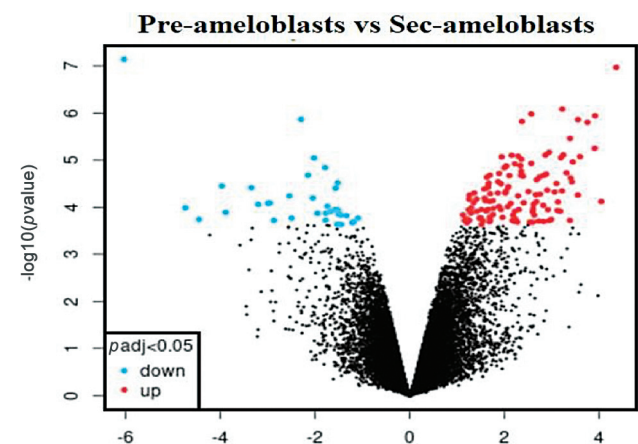


Fig. 5 Volcano plot of upregulated and downregulated genes in preameloblasts versus secretory ameloblasts.

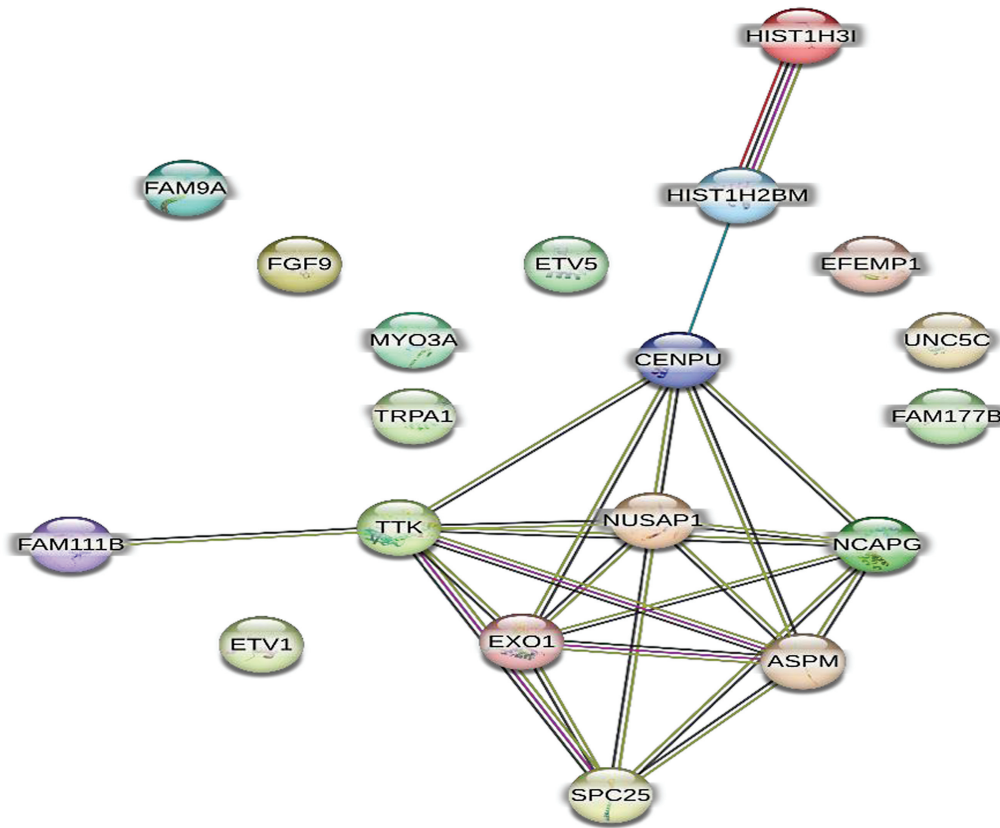


Fig. 6 Protein-protein interaction analysis.

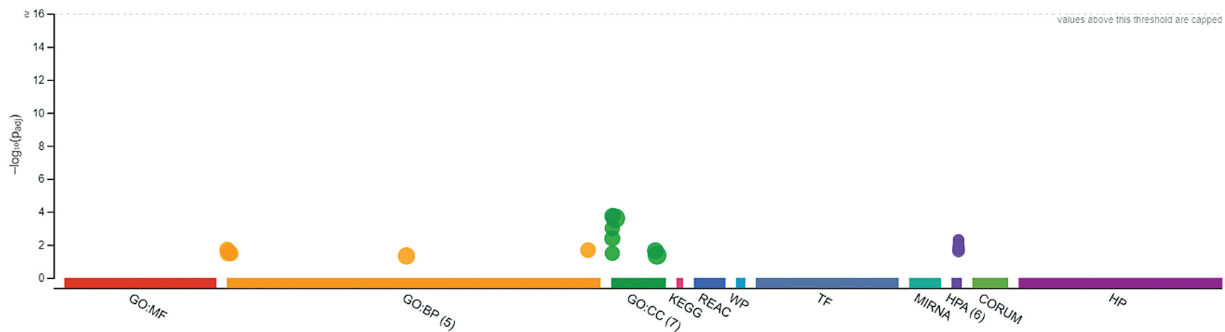


Fig. 7 Functional enrichment analysis graph.

Discussion

Ameloblasts are a type of cells responsible for the deposition of enamel and the production of extracellular matrix proteins.³ SABs deposit extracellular matrix, and the MABs mineralize the organic matrix, resulting in the production of enamel,²⁰ one of the hardest mineralized tissues in vertebrates.²¹ Three stages of ameloblast differentiation together produce a fully developed ameloblast: the precereatory, secretory, and maturation stages.² Terminally developed cells called squamous ameloblasts and molar ameloblasts are formed from undifferentiated ameloblasts called PABs during odontogenesis.²² PABs are dental epithelial cells that are elongated and columnar in shape, with

their nuclei clustered along the stellate reticulum as well as their cytoplasm, which contains the organelles required for protein synthesis and secretion of enamel. Enamel formation and ameloblast differentiation are hallmarks of the process of amelogenesis, which is governed by a number of molecular and morphogenetic events. It is generally agreed that one of the defining characteristics of amelogenesis is the change from PAB to SAB. The niche in which ameloblasts differentiate is composed of secreted proteins, enzymes, signaling molecules, and other components.

The transition from PAB to postameloblast occurs during ameloblast differentiation, and it has a profound effect on the secretion of matrix proteins in the enamel and the commencement of mineralization. Signaling molecules, their

GO:BP

Term name	Term ID	P _{adj}	$-\log_{10}(P_{adj})$	TTK	NUSAP1	UNC5C	EXO1	MYO3A	CENPU	NCAPG	FAM9A	EFEMP1	ETV5	FAM177B	FAM177A	ASPM	SPC25	ETV1	TTRN1
mitotic sister chromatid segregation	GO:0000070	1.940×10 ⁻²	1.28	■						■									
regulation of reproductive process	GO:2000241	2.096×10 ⁻²	1.31	■															
nuclear division	GO:0000280	3.062×10 ⁻²	1.51	■															
sister chromatid segregation	GO:0000819	3.287×10 ⁻²	1.54	■															
organelle fission	GO:0048285	4.746×10 ⁻²	1.68	■															

1 to 5 of 5 | Page 1 of 1

GO:CC

Term name	Term ID	P _{adj}	$-\log_{10}(P_{adj})$	TTK	NUSAP1	UNC5C	EXO1	MYO3A	CENPU	NCAPG	FAM9A	EFEMP1	ETV5	FAM177B	FAM177A	ASPM	SPC25	ETV1	TTRN1
condensed chromosome	GO:0000793	1.812×10 ⁻⁴	3.52	■															
chromosome	GO:0005694	2.463×10 ⁻⁴	3.71	■															
condensed chromosome, centromeric region	GO:0000779	1.004×10 ⁻³	2.00	■															
chromosome, centromeric region	GO:0000775	4.185×10 ⁻³	1.38	■															
chromosomal region	GO:0098687	2.286×10 ⁻²	1.64	■															
kinetochore	GO:0000776	3.254×10 ⁻²	1.54	■															
supramolecular complex	GO:0099080	4.236×10 ⁻²	1.68	■															

1 to 7 of 7 | Page 1 of 1

HPA

Term name	Term ID	P _{adj}	$-\log_{10}(P_{adj})$	TTK	NUSAP1	UNC5C	EXO1	MYO3A	CENPU	NCAPG	FAM9A	EFEMP1	ETV5	FAM177B	FAM177A	ASPM	SPC25	ETV1	TTRN1
retina; nerve fiber layer[High]	HPA:0411253	4.943×10 ⁻³	2.30	■															
retina; outer plexiform layer[High]	HPA:0411263	5.523×10 ⁻³	2.55	■															
retina; nerve fiber layer[≥Medium]	HPA:0411252	9.671×10 ⁻³	2.01	■															
retina; outer plexiform layer[≥Medium]	HPA:0411262	1.400×10 ⁻²	1.85	■															
retina; nerve fiber layer[≥Low]	HPA:0411251	1.804×10 ⁻²	1.74	■															
retina; outer plexiform layer[≥Low]	HPA:0411261	2.379×10 ⁻²	1.63	■															

1 to 6 of 6 | Page 1 of 1

Fig. 8 Table of gene ontology with terms, term identifiers (IDs), and *p*-values.

receptors, and transcriptional regulators all work together to control this process.²² However, the transcriptional pathways governing ameloblast development remain poorly understood. Specifically, we sought to identify genes that upregulate during the ameloblast transition from PAB to SAB. According to a study by Liu et al,²³ four genes (AMBN, AMTN, ENAM, and MMP-20) are associated with SABs, consistent with previous studies. They revealed that cell cycle control, deoxyribonucleic acid (DNA) damage repair, and apoptosis were linked to high-expression genes in PABs among other differentially expressed genes. On the other hand, highly expressed genes in SABs were linked to cell adhesion and extracellular matrix.²³ Another study revealed that SABs released extracellular matrix proteins into the enamel matrix consisting of three structural proteins (AMELX, AMBN, and ENAM). MMP-20 and KLK4 separate these proteins from the matrix in phases during the ameloblast secretory and maturation processes, respectively.²²

Real-time polymerase chain reaction (PCR) validated AMELX enrichment in SABs, whereas microarray analysis did not. Microarrays have a lower detection limit as opposed to real-time PCR; thus, the fold changes seen by microarrays are typically smaller.²⁴ Additional gene enrichment analysis of SABs revealed the presence of five genes (DSPP, DMP1, PHEX, ALPL, and MMP-16) involved in biomineralized tissue production, all of which have the potential to control ameloblast differentiation and enamel formation.

In this study, an analysis of gene enrichment was performed in order to conduct a comprehensive study of the

genes. In the gene enrichment study, it was discovered that of the selected genes, TTK, NUSAP1, HIST1H2BM, NCAPG, CENPU, FAM9A, SPC25, ETV5, ETV1, MYO3A, EFEMP1, and HIST1H3I are engaged in biological processes such as regulation of reproductive processes, nuclear division, and mitotic sister chromatid segregation. They are also present in cellular components, condensed chromosomes, centromeric regions, kinetochores, supramolecular complexes, and chromosomal regions. When compared to SABs, PABs have higher expression data. The genes TTK, NUSAP1, CENPU, NCAPG, FAM9A, ASPM, SPC25, and HIST1H3I function in mitotic sister chromatid segregation, cell division, and organelle fission. In comparisons of conditions of PABs and SABs, the *p*-values of these genes were lower than 0.05.

Conclusion

We compared two groups, namely, PABs and SABs. During tooth development, the expression of genes involved in the regulation of the cell cycle, repair of DNA damage, and apoptosis pathways regulate PAB maturation. Several signaling mechanisms govern SAB cell behavior, specifically the production of enamel matrix proteins and cell adhesion, both of which are required for enamel development and cell-cell interactions. PAB versus SAB expression data was used to study gene function. The TTK, NUSAP1, CENPU, NCAPG, FAM9A, ASPM, SPC25, and HIST1H3I genes function in cell division, and these genes may play a role in ameloblast development. These results will be useful in developing new methods to stimulate

ameloblast development, which is essential for tooth regeneration and tissue engineering. However, further research is required to validate the functions of these genes and the genes with which they interact. A wide variety of genetic, epigenetic, and exogenous signaling factors regulate these genes and pathways to direct throughout development and differentiation, cell fate, and behavior.

Conflict of Interest

None declared.

Acknowledgment

The author was thankful to the Deanship of Graduate Studies and Scientific Research at Dar Al Uloom University for the Support of this project.

References

- Matalová E, Lungová V, Sharpe P. Chapter 26 - development of tooth and associated structures. In: *Stem Cell Biology and Tissue Engineering in Dental Sciences*. Academic Press; 2015:335–346
- Zeichner-David M, Diekwisch T, Fincham A, et al. Control of ameloblast differentiation. *Int J Dev Biol* 1995;39(01):69–92
- Bartlett JD. Dental enamel development: proteinases and their enamel matrix substrates. *ISRN Dent* 2013;2013:684607
- Wright JT. The molecular etiologies and associated phenotypes of amelogenesis imperfecta. *Am J Med Genet A* 2006;140(23):2547–2555
- Sierant ML, Bartlett JD. Stress response pathways in ameloblasts: implications for amelogenesis and dental fluorosis. *Cells* 2012;1(03):631–645
- Gadhia K, McDonald S, Arkutu N, Malik K. Amelogenesis imperfecta: an introduction. *Br Dent J* 2012;212(08):377–379
- Stephanopoulos G, Garefalaki M-E, Lyroudia K. Genes and related proteins involved in amelogenesis imperfecta. *J Dent Res* 2005;84(12):1117–1126
- Kim J-W, Seymen F, Lin BP, et al. ENAM mutations in autosomal-dominant amelogenesis imperfecta. *J Dent Res* 2005;84(03):278–282
- Komatsu N, Takata M, Otsuki N, et al. Expression and localization of tissue kallikrein mRNAs in human epidermis and appendages. *J Invest Dermatol* 2003;121(03):542–549
- Davis S, Meltzer PS. GEO query: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* 2007;23(14):1846–1847
- Smyth GK. Limma: linear models for microarray data. In: *Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W, eds. Bioinformatics and Computational Biology Solutions using R and Bioconductor*. New York: Springer; 2005:397–420
- Smyth GK, Speed T. Normalization of cDNA microarray data. *Methods* 2003;31(04):265–273
- Murrell P. R Graphics. 1st ed. 2005Chapman and Hall/CRCNew York. <https://doi.org/10.1201/9781420035025>
- McInnes L, Healy J, Melville J. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv 2018:1802.03426
- Nuzzo RL. Histograms: a useful data analysis visualization. *PM R* 2019;11(03):309–312
- Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb)* 2015;25(02):141–151
- Li W. Volcano plots in analyzing differential expressions with mRNA microarrays. *J Bioinform Comput Biol* 2012;10(06):1231003
- Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;47(D1):D607–D613
- Raudvere U, Kolberg L, Kuzmin I, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* 2019;47(W1):W191–W198
- Bei M. Molecular genetics of ameloblast cell lineage. *J Exp Zool B Mol Dev Evol* 2009;312B(05):437–444
- Wald T, Osickova A, Sulc M, et al. Intrinsically disordered enamel matrix protein ameloblastin forms ribbon-like supramolecular structures via an N-terminal segment encoded by exon 5. *J Biol Chem* 2013;288(31):22333–22345
- Gibson CW. The Amelogenin proteins and enamel development in humans and mice. *J Oral Biosci* 2011;53(03):248–256
- Liu C, Niu Y, Zhou X, et al. Cell cycle control, DNA damage repair, and apoptosis-related pathways control pre-ameloblasts differentiation during tooth development. *BMC Genomics* 2015;16(01):592
- Gao H, Wang Y, Liu X, et al. Global transcriptome analysis of the heat shock response of *Shewanella oneidensis*. *J Bacteriol* 2004;186(22):7796–7803