Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Comparative Label-Free Quantitative Proteomics of Hypomineralised Second Primary Molars (HSPM) and Molar–Incisor Hypomineralisation (MIH) reveals divergent enamel protein signatures underpinning distinct pathogenic mechanisms

Aman Singh

123amansingh123@gmail.com

Banaras Hindu University

Rigzin Tamchos

All India Institute of Medical Sciences

Ashima Goyal

Post Graduate Institute of Medical Education and Research

Manoj Kumar Jaiswal

Post Graduate Institute of Medical Education and Research

Jaspreet Singh

Post Graduate Institute of Medical Education and Research

Umer Mukhtar

Post Graduate Institute of Medical Education and Research

Research Article

Keywords: Molar-Incisor Hypomineralisation, Hypomineralised Second Primary Molars, Enamel Proteomics, Label-Free Quantitative Proteomics, Amelogenesis, Differential Protein Abundance

Posted Date: May 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9587147/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

Despite sharing clinical features and an overlapping mineralisation window, HSPM and MIH differ at the molecular level in their enamel matrix composition, which remains uncharacterised. This study aimed to directly compare the enamel matrix proteomes of HSPM and MIH using label-free quantitative proteomics to identify differentially abundant proteins and elucidate condition-specific pathogenic mechanisms.

Methods

Enamel samples from 10 HSPM & MIH samples each were subjected to label-free quantitative LC–MS/MS analysis. Ninety proteins common to both groups were identified through secondary proteomic analysis. Sensitivity analysis evaluated detection frequencies and protein selectivity. Differential abundance was assessed using Welch's t-test with Benjamini–Hochberg false discovery rate (FDR) correction; significance was FDR < 0.05.

Results

Detection frequency analysis identified 78 proteins (86.7%) as Common High Confidence across both enamel types. Differential analysis identified 46 significantly abundant proteins (FDR < 0.05), of which 38 were enriched in HSPM enamel and 8 in MIH enamel. HSPM-enriched proteins indicated immune infiltration and a protease–antiprotease imbalance during primary molar amelogenesis, whereas MIH enamel was selectively enriched in epidermal cornification and desmosomal junction proteins, reflecting enamel organ epithelial disruption during permanent molar amelogenesis.

Conclusion

This is the first comparative enamel proteomics study of HSPM and MIH. Despite a largely shared protein pool, the two conditions exhibit distinct molecular phenotypes, with an immune-inflammatory signature in HSPM and an epithelial disruption signature in MIH, implicating divergent pathogenic pathways in primary versus permanent enamel hypomineralisation. These findings provide a molecular basis for developing condition-specific biomarkers and targeted preventive strategies.

Introduction

Enamel Hypomineralisation is a systemic defect characterised by a qualitative loss of enamel resulting from disturbances during amelogenesis, particularly during the initial calcification or maturation phase.[1] This condition is elicited in both primary and permanent dentitions. When elicited in permanent teeth, it is known as Molar Incisor Hypomineralisation (MIH), which usually affects primarily permanent first molars & permanent incisors, whereas in primary teeth, it is known as Hypomineralised Second Primary Molars

(HSPM), affecting chiefly primary second molars (Lygidakis et al. 2022). The global prevalence of enamel hypomineralisation is estimated at 25.3% according to a recently published systematic review and meta-analysis with an estimated prevalence of MIH at 15.5% (Ammar et al. 2025). For HSPM, the prevalence varies at the child and tooth levels globally, at 6.8% and 4.1%, respectively (McCarra et al. 2022). Hypomineralised enamel defects in both these conditions, i.e., HSPM & MIH, are very similar in terms of having clear, localised, demarcated borders, altered translucency of enamel, and varying colours of demarcated opacities. Given the shared period of enamel maturation, HSPM and MIH are speculated to be related; consequently, HSPM has been reported as a predictor of MIH (OR = 4.66) (Garot et al. 2018). Both these conditions present significant clinical challenges due to increased hypersensitivity, rapid post-eruptive enamel breakdown (PEB), and increased caries susceptibility, often leading to difficulty in achieving effective anaesthesia and long-term restorative success. Additionally, compromised enamel structure results in poor bonding and frequent restoration failure, while associated dental anxiety and behavioural issues further complicate clinical management in affected children (Lygidakis et al. 2022). Thus, enamel hypomineralisation in both primary and permanent dentitions constitutes a significant global oral health burden, affecting a substantial proportion of children worldwide and contributing to increased treatment needs, healthcare costs, and reduced quality of life. However, despite increasing recognition, the underlying pathogenesis remains poorly understood.

The biological basis of these conditions (HSPM & MIH) involves disruption during the initial calcification and maturation phase of amelogenesis, resulting in decreased mineral content and increased enamel porosity (Almuallem et al. 2018). This disruption during amelogenesis can be explained by the infiltration of serum albumin into the developing enamel matrix, which survives the highly proteolytic environment due to its strong affinity for hydroxyapatite crystals, thereby inhibiting their growth and leading to the formation of malformed enamel that lacks sufficient mineral content and has a higher organic portion, as seen in HSPM (Singh et al. 2025) & MIH (Mukhtar et al. 2022). In addition to the role of serum albumin, increased defective protease activity and elevated protease inhibitor levels have been proposed to contribute to this biological disruption in HSPM-affected enamel (Singh et al. 2025). The increase in the organic portion in HSPM affected enamel is contributed chiefly by proteins such as protein LEG 1 homolog, Odontogenesis-associated phosphoprotein (ODAPH), Olfactomedin like protein 3 (OFML-3), calcium-binding proteins & immunoglobins thereby suggesting an underlying infectious/inflammatory aetiology (Singh et al. 2025). In MIH-affected enamel, this contribution is dominated by collagens, Kallikrien-4 (KLK4), matrix metalloprotease-20 (MMP-20), alpha-2-HS-glycoprotein, alpha-2-macroglobulin & α1-anti-trypsin (Mukhtar et al. 2022). Thus, this significant increase in organic content in both these types of affected enamel has led to an increased structural & chemical alteration as depicted by their enamel characteristics - low mineral density and increased porosity resulting in compromised mechanical properties (Jähn-Rickert et al. 2026), (Alifakioti et al. 2021).

Proteomics is a modern, powerful, and innovative tool for identifying proteins that differ between two study groups while also offering insights into biological processes and changes in signalling pathways. Enamel proteomics mainly involves identifying proteins in the enamel by grinding the sample, followed by enzymatic digestion of proteins into smaller peptides. These peptides are then separated by liquid chromatography (LC) and further analyzed by mass spectrometry (MS). Liquid chromatography tandem mass spectrometry

(LC-MS/MS) is typically favoured among mass spectrometric techniques due to its high sensitivity, versatility, and capacity to analyse complex protein mixtures, outperforming other methods like gas chromatography with tandem mass spectrometry (GC-MS/MS) or MALDI-TOF MS. LC-MS/MS has provided valuable information regarding normal enamel compared to enamel affected by HSPM (Singh et al. 2025) & MIH (Mukhtar et al. 2022), with extended applications in neonatal teeth (Singh A et al. 2025), human saliva (Bekes et al.2020) & dental pulp (Horsophonphong et al. 2025).

Despite increasing evidence regarding the structural, clinical, and molecular features of HSPM & MIH, the existing literature mainly compares them individually with sound enamel. Although recent proteomic studies have shed light on altered protein expression and biological pathways in these conditions, there remains a lack of direct comparative analyses between HSPM and MIH-affected enamel. As a result, it remains uncertain whether these defects share a common molecular aetiology or represent distinct biological entities across the primary and permanent dentitions. This significant gap hinders our understanding of disease progression, early detection, and the development of targeted preventive and therapeutic strategies. Therefore, the present study aimed to perform a quantitative comparative proteomic analysis of HSPM and MIH-affected enamel to identify differentially expressed proteins and associated biological pathways.

Materials and Methods

This study is a secondary analysis of previously generated proteomic datasets from independent investigations of hypomineralised enamel in hypomineralised second primary molars (HSPM) (Singh et al. 2025) and molarincisor hypomineralisation (MIH) (Mukhtar et al. 2022). Ethical approval for the original studies was obtained from the Institute Ethical Committee (approval numbers: [IEC INT/2022/MDS-739] and [NK/5741/MDS/270]), and all procedures were conducted in accordance with the Declaration of Helsinki. The present study utilised de-identified raw mass spectrometry data retrieved from publicly accessible proteomic repositories- ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier (PXD056623), and no new human participants or biological samples were involved. Therefore, additional ethical approval was not required for this secondary data analysis.

Brief description of sample collection

The hypomineralised enamel of second primary molars and first permanent molars, diagnosed using the EAPD criteria (Lygidakis et al. 2022), was obtained from previously extracted & appropriately stored teeth of patients who visited the outpatient department of the Institute. For the molar incisor hypomineralisation group, only first permanent molars were included, in accordance with the methodology described by Mukhtar et al. (2022). HSPM & MIH-affected teeth indicated for extraction were duly extracted, with prior informed consent and assent from both the patients and their parents.

Brief description of sample preparation

The protocol for HSPM and MIH-affected enamel sample preparation for LC-MS/MS analysis has already been described by Singh et al. (2025) (Singh et al. 2025) and Mukhtar et al. (2022) (Mukhtar et al. 2022). Briefly, HSPM-affected & MIH-affected enamel blocks measuring 2 mm x 2 mm x 0.5 mm, were sectioned

from the individual teeth with the help of a straight diamond disc (C01/190/0.3 mm) mounted on a straight handpiece (low speed, continuous normal saline irrigation). The prepared enamel blocks were rinsed with 3% H_2O_2 for 5 minutes, washed with LC–MS-grade water, and subjected to sequential acid etching with 10% HCl, followed by a second etching step in the presence of a protease inhibitor. The extracted protein solution was desalted with a C18 resin column and eluted using acetonitrile-based solutions. The eluate was vacuum-dried and reconstituted in an ammonium bicarbonate buffer containing acetonitrile. Samples were reduced with dithiothreitol, alkylated with iodoacetamide, and digested with trypsin at 37°C for 22 hours. The reaction was terminated with formic acid, and peptides were further purified using C18 resin, vacuum-dried, and finally reconstituted in 2% acetonitrile with 0.1% formic acid for subsequent LC–MS analysis.

LC-MS/MS analysis

LC–MS/MS analysis was performed using an Orbitrap Fusion™ Tribrid™ mass spectrometer coupled with an EASY-nLC™ 1200 system via nano-electrospray ionisation. Peptides were separated on a reverse-phase nano-LC column (75 µm × 25 cm) with a 5–40% acetonitrile gradient in 0.1% formic acid over 180 minutes at a flow rate of 300 nL/min. MS1 scans were acquired in the Orbitrap at a resolution of 120,000 across a mass range of 375–1,600 m/z. Data-dependent acquisition was employed, selecting precursor ions with charge states 2–5 for MS/MS analysis. Fragmentation used higher-energy collision-induced dissociation at 30% collision energy, and MS/MS spectra were collected in the ion trap. Raw data were processed with Proteome Discoverer (v2.4) against the UniProt reference proteome database, with a precursor mass tolerance of 10 ppm and a fragment mass tolerance of 0.05 Da, allowing up to two missed tryptic cleavages. Carbamidomethylation of cysteine was set as a fixed modification, while oxidation (M) and N-terminal acetylation were considered variable modifications. The dataset was deposited in the ProteomeXchange Consortium via the PRIDE repository (identifier: PXD060636).

Proteomic Data Acquisition & Secondary LC-MS analysis

Secondary proteomic analysis was conducted using de-identified raw mass spectrometry datasets for hypomineralised second primary molars (HSPM) and molar incisor hypomineralisation (MIH) obtained from publicly accessible proteomic repositories (ProteomeXchange/PRIDE; accession numbers: [PXD056623]. The raw files were reprocessed to ensure consistency and comparability between the two groups. All datasets were processed with identical LC–MS/MS parameters using Proteome Discoverer (v2.4), with database searching against the UniProt reference proteome. Search parameters included a precursor mass tolerance of 10 ppm and a fragment mass tolerance of 0.05 Da, allowing up to two missed cleavages with trypsin. Carbamidomethylation of cysteine was set as a fixed modification, while oxidation (methionine) and N-terminal acetylation were considered variable modifications.

To reduce technical variability and enable reliable comparison, identical filtering criteria, peptide-spectrum matching thresholds, and protein identification parameters were used for both datasets. Batch effects were minimised by reprocessing all raw files using a unified analytical pipeline, ensuring consistent protein identification and quantification across datasets. The resulting normalised protein abundance data were then used for comparative quantitative proteomic analysis between HSPM- and MIH-affected enamel to identify differentially expressed proteins and related biological pathways.

Bioinformatic analysis

a) Selection of Common Proteins

To facilitate cross-dataset comparison, proteins were matched between the HSPM and MIH using shared UniProt accession numbers. Only proteins present in both datasets were kept for analysis. After removing duplicates and ambiguous entries, a final set of shared proteins was formed and used for all subsequent analyses.

b) Principal Component Analysis (PCA)

Principal component analysis (PCA) (Tang et al. 2023) was carried out on normalized protein abundance values to evaluate global proteomic differences between HSPM and MIH samples. Prior to PCA, missing values were filled with protein-wise mean substitution, and the data were standardised using z-score normalisation. PCA was performed at the sample level, with proteins as variables and samples as observations. The first two principal components (PC1 and PC2) were used to visualise clustering and separation between the HSPM and MIH groups.

c) Quantification and Data Transformation

An equal number of samples from each group (HSPM & MIH) were selected after PCA analysis for further quantification. In both groups, mean normalised abundance values for each protein were calculated across all samples within the respective group (HSPM and MIH). Protein abundance values were then log2-transformed to stabilise variance and reduce skewness inherent to proteomic intensity data.

The differential abundance of proteins between groups was calculated as:

$$\Delta \log_2 = \log_2 (\mathrm{HSPM}) - \log_2 (\mathrm{MIH})$$

This metric served as the main quantitative measure of relative protein enrichment between hypomineralised primary and permanent enamel.

d) **Scatter Plot Analysis**

To assess protein-level differences, a scatter plot (Tang et al. 2023) was generated using:

- **X-axis**: log2-transformed mean abundance in HSPM
- **Y-axis**: log2-transformed mean abundance in MIH

An identity line (y = x) was included to indicate equal protein abundance between groups. Proteins that deviated from the identity line were considered differentially abundant and contributed to the group separation observed in PCA.

e) Protein Stratification

Proteins were classified into three categories based on Δlog2 thresholds:

- **HSPM-enriched proteins**: $\Delta\log2 \geq +1$
- **MIH-enriched proteins**: $\Delta\log2 \leq -1$
- **Shared core proteins**: $-1 < \Delta\log2 < +1$

These thresholds correspond to approximately ≥ 2-fold differences in abundance and were used to detect biologically significant variations between groups.

f) STRING Network and Pathway Enrichment Analysis

Proteins identified as HSPM-enriched and MIH-enriched were analysed separately using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING©) database to evaluate protein–protein interaction (PPI) networks and functional enrichment. (https://string-db.org)

Enrichment analyses included:

- Gene Ontology (GO) Biological Process
- Molecular Function
- Cellular Component
- Reactome and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways

Statistical significance was determined using false discovery rate (FDR) adjusted p-values, while network significance was assessed based on the protein–protein interaction (PPI) enrichment p-value obtained from STRING, with a significance threshold set at ($p < 0.05$).

g) Selective Sample (Sensitivity) Analysis

To evaluate the analytical sensitivity of the LC–MS/MS workflow and to assess the consistency of enamel protein detection across individual specimens within each group, a selective sample analysis was performed on the complete set of 90 proteins identified in the shared enamel proteome of HSPM and MIH. For each protein, detection frequency was calculated independently for the HSPM group (n = 6 specimens: HSPM 2, 3, 5, 6, 7, 10) and the MIH group (n = 6 specimens: MIH 5, 6, 7, 8, 9, 10). A protein was defined as detected in a given specimen if its scaled spectral abundance value, as exported from Proteome Discoverer, was greater than zero. Detection frequency for each group was expressed as a percentage of the total specimens within that group:

Detection Frequency (%) = (Number of specimens in which protein was detected / Total specimens in group) × 100

Per-specimen detection rates were additionally calculated for each of the 12 individual enamel specimens as the percentage of the total 90-protein pool detected in that specimen, providing a measure of within-group reproducibility and specimen-level analytical sensitivity. All sensitivity analyses were performed using Python 3.12 (NumPy 1.26, Pandas 2.1). Results were visualised as (i) a detection frequency bubble scatter plot, with bubble size proportional to the absolute $\log_2$ fold change between groups; (ii) a horizontal bar chart of protein counts per selectivity class; and (iii) a per-specimen detection rate bar chart.

h. Differential abundance & Fold-change analysis

Scaled protein abundance values exported from Proteome Discoverer were used for differential analysis, with missing values imputed as zero. Data were $\log_2$-transformed using a pseudocount [$\log_2$(abundance + 1)] to stabilise variance. Mean $\log_2$ abundance values were calculated for each protein across the selected HSPM and MIH samples, and differential abundance was expressed as $\log_2$ fold change ($\log_2$FC = $\mu\log_2$(HSPM) − $\mu\log_2$(MIH)), with corresponding linear fold changes derived as 2^($\log_2$FC). Statistical significance was assessed using Welch's t-test (unequal variances) on the $\log_2$-transformed values, followed by Benjamini–Hochberg false discovery rate (FDR) correction, with significance set at FDR < 0.05. All analyses were performed in Python using SciPy and statsmodels. Results were visualised using volcano plots, ranked bar charts, and differential abundance dot plots to represent statistical significance, fold-change magnitude, and the direction of protein enrichment.

h) Statistical Considerations

All analyses utilised standard statistical and computational methods for proteomic data. To reduce technical variability, log2 transformation and z-score normalisation were performed. No imputation beyond mean substitution was applied for differential abundance analysis. Protein classification thresholds were set in advance to prevent post hoc bias.

Results

1. Identification of shared proteins between MIH & HSPM-affected Enamel

A total of 20 enamel samples were analysed, comprising 10 hypomineralised second primary molars (HSPM) and 10 molar incisor hypomineralisation (MIH)-affected permanent first molars. Proteomic profiling of all samples identified a comprehensive protein pool representing both groups. Table 1 displays a summary of the enamel protein pool identified across all 10 samples of the HSPM & MIH group. The combined dataset yielded an aggregated list of 435 proteins that constitute the overall enamel proteome, regardless of group allocation. (**Supplementary Table S1**)

Table 1
Summary of the Enamel Matrix Protein Pool Identified Across HSPM and MIH groups by Label-Free Quantitative LC–MS/MS

Category	n (proteins)	% of Relevant Total	Supplementary Reference
A. Overall Protein Pool			
Total proteins identified (HSPM & MIH)	435	—	*Supplementary Table S1*
Proteins identified in HSPM	326	75.0%	—
Proteins identified in MIH	229	52.6%	—
B. Shared and Group-Specific Proteins			
Proteins common to both HSPM and MIH	120	27.6%	*Supplementary Table S4*
Proteins unique to HSPM	206	47.4%	*Supplementary Table S2*
(% of HSPM total)	206	63.2%	—
Proteins unique to MIH	109	25.1%	*Supplementary Table S3*
(% of MIH total)	109	47.6%	—
C. Quantitative Analysis of Common Protein Pool (n = 90)			
Proteins common to both groups (for analysis)	90	100%	*Supplementary Table S4*
Common High Confidence (≥ 80% detection, both groups)	78	86.7%	—
Common Moderate (≥ 50% detection, both groups)	12	13.3%	—
D. Differential Abundance (FDR < 0.05)			
Significantly differentially abundant proteins	46	51.1%	*Supplementary Table S5*

Of the 120 proteins common to both groups, 90 were selected for quantitative differential abundance analysis based on detection in ≥ 5/10 specimens across both groups. The remaining 30 proteins (detected in < 50% of specimens in either group) were excluded from the differential analysis to ensure statistical reliability.

% of Relevant Total: For Section A, % is calculated relative to the total protein pool (n = 435). For Section B sub-rows in italics, % is calculated relative to the respective group total (HSPM: n = 326; MIH: n = 229). For Sections C and D, % is calculated relative to the 90-protein common analysis pool.

FDR: Benjamini–Hochberg false discovery rate. Significance threshold: FDR < 0.05. Statistical test: Welch two-sample t-test per protein.

Category	n (proteins)	% of Relevant Total	Supplementary Reference
A. Overall Protein Pool			
Enriched in HSPM-affected enamel	38	82.6%	—
Enriched in MIH-affected enamel	8	17.4%	—
Not significantly different (FDR ≥ 0.05)	44	48.9%	—
Total unique proteins identified across both conditions	435	100%	*Supplementary Table S1*
Of the 120 proteins common to both groups, 90 were selected for quantitative differential abundance analysis based on detection in ≥ 5/10 specimens across both groups. The remaining 30 proteins (detected in < 50% of specimens in either group) were excluded from the differential analysis to ensure statistical reliability.			
% of Relevant Total: For Section A, % is calculated relative to the total protein pool (n = 435). For Section B sub-rows in italics, % is calculated relative to the respective group total (HSPM: n = 326; MIH: n = 229). For Sections C and D, % is calculated relative to the 90-protein common analysis pool.			
FDR: Benjamini–Hochberg false discovery rate. Significance threshold: FDR < 0.05. Statistical test: Welch two-sample t-test per protein.			

Comparative analysis revealed both distinct and overlapping protein expression patterns across the two groups. Proteins were categorised into three subsets: (i) proteins uniquely expressed in the HSPM group (206 proteins) (**Supplementary Table S2)**, (ii) proteins uniquely expressed in the MIH group (109 proteins) (**Supplementary Table S3)**, and (iii) proteins common to both groups (120 proteins) (**Supplementary Table S4)**. The common protein subset represents the shared molecular framework of hypomineralised enamel, whereas the unique subsets reflect group-specific biological variations in enamel development, maturation, or pathological processes.

This structured dataset served as the basis for downstream functional annotation, including classification into molecular and cellular functions and into broader functional categories such as structural proteins, enzymes, chaperones, extracellular matrix components, transport proteins, and protease inhibitors.

Among the 120 proteins identified as common to both MIH and HSPM groups, a subset was further refined for comparative analysis. Specifically, only proteins detected in ≥ 50% of samples (i.e., present in at least 5 out of 10 samples) within each group were considered eligible. Based on this criterion, 90 proteins met the inclusion threshold and were retained for downstream comparative and functional analyses. This filtering approach ensured the selection of consistently expressed proteins, thereby enhancing the reliability and biological relevance of the comparative proteomic evaluation (**Supplementary Table S4).**

2. Global Proteomic Differences Between HSPM and MIH

Principal Component Analysis (PCA)

Principal component analysis revealed distinct yet partially overlapping clusters of HSPM and MIH samples (Fig. 1). The first principal component (PC1) accounted for most of the variance and primarily separated HSPM from MIH samples, whereas the second principal component (PC2) captured intra-group variability. PCA of the data is shown in Fig. 2, which includes 10 samples from both the MIH and HSPM groups. Samples labelled MIH-5, MIH-6, MIH-7, MIH-8, MIH-9, and MIH-10 from the MIH group, and those labelled HSPM-2, HSPM-3, HSPM-5, HSPM-6, HSPM-7, and HSPM-10 from the HSPM group, were closely related, with minimal variability. HSPM samples exhibited relatively compact clustering, suggesting a more homogeneous proteomic profile, whereas MIH samples showed greater dispersion, indicative of increased heterogeneity in permanent enamel hypomineralisation. These findings indicate that although HSPM and MIH share underlying proteomic features, they represent distinct biological phenotypes. Consequently, further analysis focused on these specific samples from both groups.

3. Protein-level differences: Scatter Plot analysis

Scatter plot analysis of log2-transformed protein abundances revealed systematic differences in protein distributions between HSPM and MIH enamel. Proteins clustering along the identity line (y = x) showed comparable abundance in both groups, whereas those deviating from the diagonal showed differential enrichment and contributed to the separation observed in PCA. Proteins enriched in MIH enamel showed greater deviation from the identity line, indicating larger abundance differences across individual proteins, whereas a comparable or greater number of proteins showed enrichment in HSPM, albeit with a more modest deviation from the diagonal. (**Supplementary Fig. 1**)

4. Proteins Stratification based on differential normalised abundances

Using predefined Δlog2 thresholds, proteins were categorised into three groups:

- • HSPM-enriched proteins: $\Delta\log2 \geq +1$ (n = 52)
- • MIH-enriched proteins: $\Delta\log2 \leq -1$ (n = 26)
- • Shared core proteins: $-1 < \Delta\log2 < +1$ (n = 12)

The predominance of HSPM-enriched proteins suggests a marked shift in the proteomic composition of primary enamel compared with permanent enamel. (**Supplementary Table S5**)

4.1 HSPM-Enriched Proteins

Proteins enriched in the HSPM group were predominantly involved in extracellular matrix organisation, structural integrity, and regulated biomineralisation. Many of these proteins were associated with cell adhesion, matrix assembly, and calcium-binding activity, reflecting a microenvironment supportive of organised enamel formation. The presence of proteins involved in ion transport, protein folding, and stabilisation further suggests ongoing modulation of mineral deposition and maintenance of matrix architecture. Collectively, the HSPM-enriched proteomic profile indicates a relatively coordinated and

adaptive biological environment, characterised by structural preservation and controlled mineralisation dynamics despite hypomineralisation.

4.2 MIH-Enriched Proteins

Proteins enriched in the MIH group were largely associated with stress-related pathways, proteolytic activity, and dysregulated cellular processes. A significant proportion of these proteins was associated with oxidative stress responses, inflammation, and protein degradation mechanisms, indicating a compromised enamel developmental milieu. Additionally, proteins involved in metabolic imbalance and impaired protein processing were observed, suggesting altered cellular homeostasis and increased susceptibility to structural breakdown. This proteomic pattern reflects a predominantly stress-driven, degradative environment, consistent with pathological remodelling and defective mineralisation characteristic of MIH.

4.3 Shared Core Proteins

The shared core proteins showed relatively stable and comparable expression across the HSPM and MIH groups, underscoring their essential role in fundamental enamel biology. These proteins were primarily involved in housekeeping functions, including basic cellular metabolism, maintenance of structural components, and essential protein-processing pathways. Their consistent abundance across both conditions suggests that they form a conserved proteomic backbone necessary for enamel formation and maintenance, and that they remain largely unaffected by the distinct pathological processes that influence mineralisation in HSPM and MIH.

5. **Protein–Protein Interaction and Pathway Enrichment Analysis**

Proteins showing enrichment in the respective groups, i.e., MIH, HSPM or shared core group, were uploaded in the **STRING©** (Search Tool for the Retrieval of Interacting Genes/Proteins) software to obtain gene ontology annotations and **KEGG** (Kyoto Encyclopedia of Genes and Genomes) pathways. STRING network analysis revealed significant protein–protein interaction enrichment in both the HSPM- and MIH-enriched protein sets, suggesting that these proteins are functionally interconnected rather than randomly distributed. Both KEGG pathway and Local network clustering showed enrichment for complement and coagulation cascade and positive regulation of opsonisation.

5.1 HSPM-Enriched Network

HSPM-enriched proteins demonstrated significant functional enrichment across multiple biological processes, prominently including the acute-phase inflammatory response, regulation and negative regulation of proteolysis, and enzymatic activities such as endopeptidase and hydrolase functions. Additionally, pathways related to the positive regulation of receptor-mediated endocytosis were enriched, with signal intensities gradually decreasing across these processes, as illustrated in **Supplementary Fig. 2**. Reactome pathway analysis further revealed strong enrichment for pathways associated with platelet and neutrophil degranulation, as well as platelet activation, signalling, and aggregation as depicted in Fig. 2A. These findings indicate a coordinated involvement of inflammatory modulation, controlled proteolytic activity, and vesicular transport mechanisms within the HSPM proteomic profile.

5.2 MIH-Enriched Network

In contrast, MIH-enriched proteins showed significant functional enrichment in biological processes related to keratinocyte differentiation, intermediate filament organisation, keratinisation, and broader cellular and cytoskeletal organisation as illustrated in **Supplementary Fig. 3**. Correspondingly, molecular function and cellular component analyses revealed enrichment for structural constituents of the cytoskeleton, with localisation to keratin, intermediate filaments, and the cornified envelope. Similar patterns of functional enrichment were observed within local network clusters, reinforcing the dominance of cytoskeleton-associated pathways. KEGG pathway analysis showed significant enrichment for pathways associated with Staphylococcus aureus infection, suggesting potential links to host defence and inflammatory responses. Reactome pathway analysis further highlighted enrichment in cornified envelope formation, apoptosis, oxidative stress–induced senescence, and innate immune system pathways as illustrated in Fig. 2B. Collectively, these findings indicate a MIH proteomic profile characterised by altered epithelial differentiation, cytoskeletal remodelling, and stress and immune-associated biological processes.

5.3 Shared Core Enrichment

The shared core protein group showed limited but consistent enrichment of fundamental cellular and metabolic processes, indicating a conserved proteomic framework essential for enamel formation and maintenance. However, compared with the HSPM- and MIH-enriched groups, these proteins exhibited lower pathway specificity, supporting their role as a stable baseline proteome rather than as drivers of condition-specific biological variation.

6. Differential Protein Abundance analysis between MIH and HSPM

6A. Volcano plot analysis

Volcano plot analysis (Fig. 3A**)** of the enamel matrix proteome revealed that 46 of the 90 proteins common to both HSPM and MIH enamel were significantly differentially abundant (FDR < 0.05), with 38 enriched in HSPM-affected enamel and 8 in MIH-affected enamel. The most markedly HSPM-enriched enamel proteins were the immunoglobulin *IGHG1* (log_2FC = + 7.46; 175.7-fold), fibrillar collagen *COL1A1* (+ 6.96; 124.5-fold), and the plasma transport proteins *APOA1* (+ 6.72; 105.4-fold), *ORM1* (+ 6.70; 103.7-fold), and *VTN* (+ 6.69; 103.6-fold), alongside *SERPINF1* (+ 6.00; 64.1-fold) and the enamel matrix protease *MMP20* (+ 5.72; 52.6-fold). In contrast, the MIH-enriched enamel proteins were characterised by the histone variant *H2AC20* (log_2FC = − 6.41; 84.7-fold higher in MIH), the antimicrobial peptide *DCD* (− 6.37; 82.5-fold), the cornification protein *HRNR* (− 5.33; 40.3-fold), the calcium-binding protein *ANXA1* (− 5.30; 39.4-fold), and the epidermal structural protein *KPRP* (− 4.98; 31.6-fold). The 44 non-significant proteins clustered below the FDR threshold irrespective of fold-change magnitude, reflecting comparable abundance in the enamel matrix of both hypomineralised dentition types.

6B. Ranked Bar Chart for differentially abundant significant proteins between MIH & HSPM

The **ranked bar chart (**Fig. 3B**)** offers a comprehensive overview of all 46 differentially abundant proteins. Among HSPM-enriched proteins, the highest log_2FC values were observed for immunoglobulins (IGHG1, + 7.46; IGHV4-34, + 6.40; IGKV3-20, + 5.51; IGHG4, + 4.56; IGKC, + 4.43), plasma transport proteins (APOA1, + 6.72; GC, + 5.11; TTR, + 5.10; TF, + 4.06), ECM and collagen proteins (COL1A1, + 6.96; COL11A2, + 6.18; COL17A1, + 3.17; COL1A2, + 2.97), and serine protease inhibitors (SERPINF1, + 6.00; SERPINA1, + 5.77; SERPINC1, + 4.17; SERPINA3, + 2.90). Among the eight MIH-enriched proteins, cornification and epidermal structural proteins predominated: H2AC20 (− 6.41), DCD (− 6.37), HRNR (− 5.33), ANXA1 (− 5.30), KPRP (− 4.98), KRT9 (− 3.58), DSG1 (− 3.03), and DSP (− 2.88), all validated at FDR < 0.05.

6C. Differential abundance dot plot analysis & top-hit summary

Figure 3C illustrates differential abundance dot plot provides an integrated view of fold-change magnitude and statistical confidence for all 46 significant enamel proteins. Proteins in the upper-right quadrant, characterised by high log_{10}(p-value) and large positive log_2FC, are the most statistically robust and quantitatively dominant in HSPM-affected enamel, including COL1A1, APOA1, ORM1, VTN, and MMP20. Their large dot sizes reflect extreme linear fold-change magnitudes (52–125×) in HSPM relative to MIH enamel. The upper-left quadrant, occupied by H2AC20, DCD, KPRP, and HRNR, identifies the most statistically significant and abundantly enriched proteins in MIH-affected permanent enamel. The top-hit summary reinforces the pronounced quantitative asymmetry between the two conditions: HSPM-enriched proteins reach dramatically higher maximum fold changes (up to 175.7× for IGHG1) than MIH-enriched proteins (up to 84.7× for H2AC20). This asymmetry likely reflects the more extensive incorporation of immune and systemic proteins within the primary enamel matrix of HSPM, where the ameloblast secretory window coincides with an immunologically vulnerable prenatal-to-early-postnatal period, compared with the more localised epithelial disruption that characterises MIH-affected permanent enamel formed during early childhood. Table 2 summarises 46 significantly differentially abundant enamel matrix proteins with FDR < 0.05, and a complete statistical analysis for all 90 differentially abundant enamel matrix proteins is provided in **Supplementary Table S6.**

Table 2

Significantly differentially abundant enamel matrix proteins: HSPM-enriched and MIH-enriched (FDR < 0.05)

PROTEINS SIGNIFICANTLY ENRICHED IN HSPM (n = 38) — Ranked by Descending log_2FC							
No.	**UniProt Accession**	**Gene Symbol**	**Protein Name**	**log_2FC (HSPM/MIH)**	**Fold Change (×)**	**FDR**	**Sig.**
1	**P01857**	**IGHG1**	Immunoglobulin heavy constant gamma 1	**7.457**	**175.68**	0.0079	**
2	**P02452**	**COL1A1**	Collagen alpha-1(I) chain	**6.96**	**124.47**	0.0003	***
3	**P02647**	**APOA1**	Apolipoprotein A-I	**6.72**	**105.44**	0.0005	***
4	**P02763**	**ORM1**	Alpha-1-acid glycoprotein 1	**6.696**	**103.67**	0.0005	***
5	**P04004**	**VTN**	Vitronectin	**6.695**	**103.58**	0.0005	***
6	**P06331**	**IGHV4-34**	Immunoglobulin heavy variable 4–34	**6.4**	**84.46**	0.0022	**
7	**P13942**	**COL11A2**	Collagen alpha-2(XI) chain	**6.175**	**72.28**	0.0079	**
8	**P36955**	**SERPINF1**	Pigment epithelium-derived factor	**6.002**	**64.07**	0.0009	***
9	**P01009**	**SERPINA1**	Alpha-1-antitrypsin	**5.77**	**54.56**	0.0135	*
10	**O60882**	**MMP20**	Matrix metalloproteinase-20	**5.716**	**52.58**	0.0005	***
11	**P01619**	**IGKV3-20**	Immunoglobulin kappa variable 3–20	**5.509**	**45.54**	0.0021	**
12	**P06702**	**S100A9**	Protein S100-A9	**5.426**	**42.99**	0.0023	**
13	**P02774**	**GC**	Vitamin D-binding protein	**5.105**	**34.42**	0.0022	**
14	**P02766**	**TTR**	Transthyretin	**5.101**	**34.32**	0.0052	**
15	**P69905**	**HBA1**	Hemoglobin subunit alpha	**4.867**	**29.18**	0.0032	**
16	**P02765**	**AHSG**	Alpha-2-HS-glycoprotein	**4.659**	**25.26**	0.0132	*

Notes & abbreviations - log_2FC: log_2 fold change (HSPM/MIH). Positive values = higher in HSPM; negative values = higher in MIH. Fold Change (×): For HSPM-enriched proteins, values show how many times more abundant the protein is in HSPM vs. MIH. For MIH-enriched proteins, values show fold enrichment in MIH over HSPM (inverted ratio: 1/FC). FDR: Benjamini–Hochberg False Discovery Rate. Significance threshold: FDR < 0.05. Significance codes: *** FDR < 0.001 ** FDR < 0.01 * FDR < 0.05. Statistical test: Welch two-sample t-test applied per protein across 6 HSPM and 6 MIH samples.

PROTEINS SIGNIFICANTLY ENRICHED IN HSPM (n = 38) — Ranked by Descending log_2FC							
No.	**UniProt Accession**	**Gene Symbol**	**Protein Name**	**log_2FC (HSPM/MIH)**	**Fold Change (×)**	**FDR**	**Sig.**
17	**P08493**	**MGP**	Matrix Gla protein	**4.657**	**25.22**	0.0021	**
18	**P01861**	**IGHG4**	Immunoglobulin heavy constant gamma 4	**4.557**	**23.54**	0.0251	*
19	**P01834**	**IGKC**	Immunoglobulin kappa constant	**4.426**	**21.5**	0.0263	*
20	**P01008**	**SERPINC1**	Antithrombin-III	**4.171**	**18.01**	0.0189	*
21	**P68871**	**HBB**	Hemoglobin subunit beta	**4.089**	**17.02**	0.0067	**
22	**P02787**	**TF**	Serotransferrin	**4.062**	**16.7**	0.0079	**
23	**Q15113**	**PCOLCE**	Procollagen C-endopeptidase enhancer 1	**4.027**	**16.3**	0.0025	**
24	**P00734**	**F2**	Prothrombin	**3.864**	**14.56**	0.034	*
25	**P80748**	**IGLV3-21**	Immunoglobulin lambda variable 3–21	**3.696**	**12.96**	0.0041	**
26	**A0A087WW87**	**IGKV2-40**	Immunoglobulin kappa variable 2–40	**3.669**	**12.72**	0.0299	*
27	**Q9UMD9**	**COL17A1**	Collagen alpha-1(XVII) chain	**3.17**	**9**	0.0002	***
28	**P61769**	**B2M**	Beta-2-microglobulin	**3.166**	**8.97**	0.0009	***
29	**P02788**	**LTF**	Lactotransferrin	**3.094**	**8.54**	0.0288	*
30	**P10909**	**CLU**	Clusterin	**3.029**	**8.16**	0.0259	*
31	**P08123**	**COL1A2**	Collagen alpha-2(I) chain	**2.967**	**7.82**	0.0005	***
32	**P01011**	**SERPINA3**	Alpha-1-antichymotrypsin	**2.901**	**7.47**	0.0392	*

Notes & abbreviations - log_2FC: log_2 fold change (HSPM/MIH). Positive values = higher in HSPM; negative values = higher in MIH. Fold Change (×): For HSPM-enriched proteins, values show how many times more abundant the protein is in HSPM vs. MIH. For MIH-enriched proteins, values show fold enrichment in MIH over HSPM (inverted ratio: 1/FC). FDR: Benjamini–Hochberg False Discovery Rate. Significance threshold: FDR < 0.05. Significance codes: *** FDR < 0.001 ** FDR < 0.01 * FDR < 0.05. Statistical test: Welch two-sample t-test applied per protein across 6 HSPM and 6 MIH samples.

PROTEINS SIGNIFICANTLY ENRICHED IN HSPM (n = 38) — Ranked by Descending log_2FC							
No.	**UniProt Accession**	**Gene Symbol**	**Protein Name**	**log_2FC (HSPM/MIH)**	**Fold Change (×)**	**FDR**	**Sig.**
33	**P01023**	**A2M**	Alpha-2-macroglobulin	**2.785**	**6.89**	0.0488	*
34	**Q9NZW4**	**DSPP**	Dentin sialophosphoprotein	**2.696**	**6.48**	0.0022	**
35	**P61626**	**LYZ**	Lysozyme C	**2.662**	**6.33**	0.0077	**
36	**P04259**	**KRT6B**	Keratin, type II cytoskeletal 6B	**2.552**	**5.87**	0.0259	*
37	**P01024**	**C3**	Complement C3	**2.282**	**4.86**	0.0392	*
38	**P01599**	**IGKV1-17**	Immunoglobulin kappa variable 1–17	**1.37**	**2.58**	0.0392	*
PROTEINS SIGNIFICANTLY ENRICHED IN MIH (n = 8) — Ranked by Descending log_2FC							
No.	**UniProt Accession**	**Gene Symbol**	**Protein Name**	**log_2FC (HSPM/MIH)**	**Fold Change (×)**	**FDR**	**Sig.**
1	**Q16777**	**H2AC20**	Histone H2A type 2-C	**-6.405**	**84.75**	0.0022	**
2	**P81605**	**DCD**	Dermcidin	**-6.366**	**82.46**	0.003	**
3	**Q86YZ3**	**HRNR**	Hornerin	**-5.332**	**40.27**	0.0044	**
4	**P04083**	**ANXA1**	Annexin A1	**-5.302**	**39.45**	0.0189	*
5	**Q5T749**	**KPRP**	Keratinocyte proline-rich protein	**-4.981**	**31.59**	0.0032	**
6	**P35527**	**KRT9**	Keratin, type I cytoskeletal 9	**-3.578**	**11.94**	0.0077	**
7	**Q02413**	**DSG1**	Desmoglein-1	**-3.03**	**8.17**	0.0067	**
8	**P15924**	**DSP**	Desmoplakin	**-2.883**	**7.38**	0.0263	*

Notes & abbreviations - log_2FC: log_2 fold change (HSPM/MIH). Positive values = higher in HSPM; negative values = higher in MIH. Fold Change (×): For HSPM-enriched proteins, values show how many times more abundant the protein is in HSPM vs. MIH. For MIH-enriched proteins, values show fold enrichment in MIH over HSPM (inverted ratio: 1/FC). FDR: Benjamini–Hochberg False Discovery Rate. Significance threshold: FDR < 0.05. Significance codes: *** FDR < 0.001 ** FDR < 0.01 * FDR < 0.05. Statistical test: Welch two-sample t-test applied per protein across 6 HSPM and 6 MIH samples.

8. **Selective Sample (Sensitivity) Analysis**:

Figure 4 illustrates Sensitivity analysis of the HSPM and MIH enamel proteomes. To evaluate analytical sensitivity and protein selectivity across groups, detection frequency was computed for each of the 90

identified proteins across the six HSPM and six MIH samples separately. The detection-frequency scatter (Fig. 4a) showed that most proteins clustered in the high-detection quadrant (≥ 80% in both groups), with bubble sizes reflecting the magnitude of the intergroup fold change. Selectivity classification (Fig. 4b) confirmed that 86.7% of proteins (78/90) formed a high-confidence common core, with the remaining 13.3% (12/90) in the moderate-common category; no protein met the threshold for group selectivity, indicating that intergroup differences are primarily quantitative rather than qualitative. Per-sample detection rates (Fig. 4c) were uniformly high with 87–98% across HSPM samples and 100% in all MIH samples thereby validating the sensitivity and reproducibility of the label-free quantitative proteomics workflow. These findings support the dataset's robustness and reinforce the reliability of the differential abundance and enrichment analyses. **Supplementary Table 7** provides a summary of the sensitivity analysis along with protein detection frequency and per-specimen detection rates across the HSPM and MIH enamel proteome.

Discussion

Previous studies on proteomics have shed light on various specific aspects, such as Mukhtar et al. 2022 (Mukhtar et al. 2022), which examined Label free Quantitative proteomics of MIH versus sound permanent enamel; Singh et al. 2025 (Singh et al. 2025), focusing on proteomics of HSPM versus sound primary enamel; Rexhaj et al. 2024 (Rexhaj et al. 2024), conducting proteomic profiling of MIH by severity grade; and Elfrink et al. 2015 (Elfrink et al. 2015), which called for standardised HSPM/MIH studies. To our knowledge, the present study is the first to characterise and directly compare the enamel matrix proteomes of hypomineralised second primary molars (HSPM) and molarincisor hypomineralisation (MIH) using label-free quantitative LC–MS/MS, thereby bridging a significant gap in the comparative molecular pathology of primary and permanent enamel hypomineralisation. This in vitro study identified 90 proteins that were eligible for quantitative comparison and common to both the HSPM and MIH groups. Among these 90 proteins, 46 were significantly differentially abundant (FDR < 0.05), with pronounced asymmetry observed (38 HSPM-enriched vs. 8 MIH-enriched).

HSPM has also been identified as a predictor of MIH, with an odds ratio of 4.64 (Garot et al 2018), (Negre-Barber et al. 2023). This is because the mineralisation periods of second primary molars and first permanent molars overlap during birth to 1 year of age (Elfrink et al. 2024) A second factor that has contributed to this prediction is the identification of albumin which acts as biomarker as well as mineralisation inhibitor in hypomineralised enamel thereby playing a key role in this pathomechanism (Perez et al. 2020). This shared enamel proteome core was also evident in this current study as 86.7% of the enamel protein pool was common to both conditions at high detection confidence (≥ 80% in both groups), indicating that HSPM and MIH share near-identical qualitative enamel matrix protein inventories. The biological distinctions between the two conditions are primarily quantitative (abundance-level), not qualitative (presence–absence). This aligns with HSPM's established role as a clinical predictor of MIH.

Previous studies on HSPM enamel proteomics have found elevated levels of proteins such as immunoglobulins and complement, thereby implicating an underlying inflammatory or infectious aetiology. They have also highlighted key biological processes, such as complement coagulation cascade activation and neutrophil degranulation, in HSPM enamel (Singh et al.2025). Similar biological processes (neutrophil-

mediated immunity and complement activation) are also established in MIH enamel (Pappa et al. 2022). These altered biomechanics, along with albumin infiltration, lead to an increased protein load in the developing enamel, with an estimated 3–15-fold higher protein content in hypomineralised enamel compared with normal enamel (Mangum et al. 2010). In the present study, the most striking feature of the HSPM enamel proteome was the massive enrichment of immunoglobulins (IGHG1, IGHV4-34, IGKV3-20, IGHG4, IGKC, IGKV2-40, IGLV3-21, IGKV1-17), complement (C3), haemoglobin subunits (HBA1, HBB), coagulation factors (F2), and plasma transport proteins (APOA1, VTN, ORM1, GC, TTR, TF, AHSG). These proteins are not part of the normal enamel matrix; their presence indicates a compromised ameloblast barrier, allowing systemic protein infiltration during primary molar amelogenesis. Systemic infections or inflammatory episodes during the prenatal-to-early-postnatal mineralisation window (19th week of gestation to 1 year of age) may activate complement, trigger immunoglobulin production, and disrupt ameloblast tight junctions, thereby collectively allowing immune proteins to infiltrate and become entrapped in the incompletely mineralised primary enamel matrix, directly impeding hydroxyapatite crystal growth.

During amelogenesis, proteins like kallikrein-4 (KLK4) and matrix metalloproteinase 20 (MMP20) are crucial for removing enamel matrix proteins and promoting mineralisation (Simmer et al. 2002). This remodelling process, involving proteolytic cleavage and signal transduction, was emphasised by null mice phenotypes showing that the absence of MMP20 disrupts cell-cell attachments, leading to hypomineralised enamel with increased protein retention (Bartlett et al. 2013). Furthermore, Mukhtar et al. (2022) (Mukhtar et al. 2022) proposed that an imbalance between KLK4 and antithrombin-III (SERPINC1) may contribute to poor mineralisation in MIH through their proteomic analysis. Four serpins - SERPINF1, SERPINA1, SERPINC1, SERPINA3 were notably enriched in HSPM enamel proteome (log_2FC: +2.90 to + 6.00). Since serpins are potent inhibitors of serine proteases such as KLK4 and neutrophil elastase, excessive serpin activity could suppress KLK4's role in degrading enamel matrix proteins during maturation, leading to protein accumulation and incomplete mineralisation. Additionally, there were increased collagen proteins found to be persistent in the HSPM proteome, such as COL1A1 (+ 6.96; 124.5-fold), COL11A2 (+ 6.18; 72.3-fold), COL1A2 (+ 2.97; 7.8-fold), and COL17A1 (+ 3.17; 9.0-fold). These fibrillar collagens are absent in normal mature enamel. Their accumulation in HSPM primary enamel is consistent with impaired matrix metalloprotease-mediated degradation during the maturation stage, or with vascular/connective tissue invasion of the enamel organ during primary amelogenesis. The presence of mature collagen in the organic matrix of hypomineralised enamel is also confirmed in MIH enamel (Acil et al. 2005). However, Mukhtar et al. (2022) (Mukhtar et al. 2022) observed a reduction in collagen biosynthesis proteins in MIH enamel compared to sound enamel.

Unlike the HSPM proteome, the MIH-enriched proteome mainly consisted of epidermal structural and cornification proteins, including H2AC20 (H2A clustered Histone 20), DCD (dermcidin), HRNR (hornerin), ANXA1 (annexin A1), KPRP (keratinocyte proline-rich protein), KRT9 (keratin 9), DSG1 (desmoglein-1), and DSP (desmoplakin). These proteins are key components of the cornified cell envelope in stratified squamous epithelium and are not typically part of the permanent enamel matrix. Their presence in MIH enamel indicates abnormal activation of epidermal keratinisation pathways within the enamel organ epithelium during permanent molar amelogenesis. MIH is already established as a qualitative enamel defect (Weerheijm et al. 2001) occurring during the permanent first molar amelogenesis, with a multifactorial

aetiology involving gene-environment interactions (Garot et al. 2022). Since permanent molar enamel mineralisation occurs from birth to age 3–5; any disruption occurring during this period causes MIH (Elfrink et al. 2024). Thus, it can be hypothesised that the systemic disturbances during early childhood (febrile illness, antibiotic exposure, hypoxia) acting during permanent molar amelogenesis may induce stress-response keratinisation in enamel organ epithelial cells, triggering desmosomal disassembly (DSG1 loss, DSP accumulation) and cornification programmes (HRNR, KPRP, DCD incorporation), collectively disrupting ameloblast intercellular adhesion and impairing the coordinated secretory and maturation activities required for normal permanent enamel formation.

The proteins enriched in the HSPM proteome also include MMP20 (enamelysin; $log_2FC = + 5.72$; 52.6-fold), which was markedly increased despite its usual role in degrading matrix proteins during amelogenesis. This suggests a paradoxical proteomic profile. Possible explanations are: (i) impaired regulation of MMP20 catalytic activity due to excess serpins (notably SERPINA1/SERPINC1) in HSPM enamel; (ii) failure to remove MMP20 from the matrix after secretion; or (iii) secondary accumulation of inactive MMP20 in the incompletely remodelled HSPM matrix. Other significantly enriched proteins include Matrix Gla Protein (MGP) and Fetuin-A (AHSG), both of which are known inhibitors of abnormal or ectopic mineralisation. Their accumulation in primary HSPM enamel may hinder hydroxyapatite crystal formation, contributing to the incomplete mineralisation observed in HSPM opacities. This mechanism has been established for albumin and other plasma proteins, which inhibit enamel mineralisation (Perez et al. 2020), with evidence showing that aged albumin demarcates chalky regions of hypomineralised enamel (Hubbard et al. 2017). Fetuin-A (AHSG) is also recognised as a systemic inhibitor of mineralisation (Arnaud et al. 2002).

Limitations

Although this study establishes new dynamics in HSPM & MIH aetiology through label free quantitative proteomics workup, the following limitations should be acknowledged

1. **(i) Sample size**: Although 10 HSPM and 10 MIH enamel samples were compared, final analysis with 6 samples of each group may limit statistical power. Replication with larger, multicentre cohorts is needed.
2. **(ii) Enamel extraction**: Protein contamination from the enamel pellicle (including albumin) cannot be entirely excluded despite surface cleaning, as cautioned by Hubbard et al. (2017) (Hubbard et al. 2017).
3. **(iii) In vitro study**: This is an in vitro LC–MS/MS study. Findings require validation at the protein level (Western blot, immunohistochemistry) and in larger clinical cohorts.
4. **(iv) Causality**: Differential protein abundance identifies association, not causality. Longitudinal and functional studies are needed to confirm the proposed mechanistic pathways.
5. **(v) Severity grading**: HSPM and MIH specimens were not sub-stratified by clinical severity grade, which may introduce heterogeneity within groups.

Conclusion

The present study provides the first direct, comparative characterisation of the enamel matrix proteomes of hypomineralised second primary molars (HSPM) and molar incisor hypomineralisation (MIH) using label-

free quantitative LC–MS/MS. Analysis of 90 proteins common to both hypomineralised enamel types revealed a broadly shared qualitative protein pool, with 86.7% detected consistently across both groups, indicating that HSPM and MIH incorporate a near-identical complement of enamel matrix proteins at the level of binary detection. However, quantitative differential abundance analysis identified 46 proteins (51.1%) as significantly different between the two conditions (FDR < 0.05), with 38 enriched in HSPM-affected primary molar enamel and 8 in MIH-affected permanent molar enamel.

The HSPM enamel proteome was characterised by pronounced enrichment of immunoglobulins, complement factors, collagen isoforms, serine protease inhibitors, and plasma-derived transport proteins a molecular signature consistent with immune infiltration of the primary enamel matrix and protease–antiprotease imbalance during primary molar amelogenesis, arising from systemic inflammatory or infectious disturbances during the prenatal-to-early-postnatal developmental window. In contrast, the MIH enamel proteome exhibited selective enrichment of epidermal structural proteins, cornification-associated proteins, and desmosomal junction components a profile indicative of aberrant keratinisation and epithelial disruption of the enamel organ during permanent molar amelogenesis, consistent with early-childhood systemic disturbances acting during the critical mineralisation window of the first permanent molars.

Unsupervised hierarchical clustering independently validated these findings, achieving complete segregation of HSPM and MIH enamel specimens based solely on protein abundance profiles, without prior group labelling. Together, these data demonstrate that despite their clinical and histological similarities, HSPM and MIH are underpinned by distinct molecular phenotypes in their enamel matrices an immune-inflammatory pathway in primary hypomineralisation and an epithelial disruption pathway in permanent hypomineralisation. These findings advance our understanding of the differential pathogenesis of enamel hypomineralisation across dentitions and identify candidate condition-specific protein signatures that may inform future biomarker development and targeted preventive strategies.

Abbreviations

FDR
false discovery rate
HSPM
Hypomineralised Second Primary Molars
LC–MS/MS
liquid chromatography tandem mass spectrometry
MIH
Molar-Incisor Hypomineralisation.

Declarations

Conflict of Interest statement

The authors declare no potential conflict of interest with respect to research, authorship or publication of this work.

Author Contribution

AS contributed to visualisation and conceptualisation, methodology, validation, formal analysis, investigation, data curation, and writing – original draft preparation. UM assisted in preparing MIH proteomics data and contributed to the analysis and interpretation of the data. RT contributed to the interpretation of results and critically revised the manuscript. AG contributed to analysis and interpretation, writing, review and editing, and critically revised the manuscript. MK contributed to workflow, design, resources, and data curation. JS contributed to methodology, validation, and formal analysis. All the authors gave final approval and agree to be accountable for all aspects of the work.

Acknowledgement

We would like to acknowledge the CSIC (Central Sophisticated Instrument Cell) department of PGIMER Chandigarh for their support in the accomplishment of this work.

Data Availability

Secondary proteomic analysis was conducted using de-identified raw mass spectrometry datasets for Hypomineralized Second Primary Molars (HSPM) and Molar Incisor Hypomineralization (MIH) obtained from publicly accessible proteomic repositories (ProteomeXchange/PRIDE; accession numbers: [PXD056623]. The raw files were reprocessed to ensure consistency and comparability between the two groups. All datasets were processed with identical LC–MS/MS parameters using Proteome Discoverer (v2.4), with database searching against the UniProt reference proteome. Search parameters included a precursor mass tolerance of 10 ppm and a fragment mass tolerance of 0.05 Da, allowing up to two missed cleavages with trypsin. Carbamidomethylation of cysteine was set as a fixed modification, while oxidation (methionine) and N-terminal acetylation were considered variable modifications.

References

1. Acil Y, Mobasseri AE, Warnke PH, Terheyden H, Wiltfang J, Springer I. Detection of mature collagen in human dental enamel. Calcif Tissue Int. 2005;76(2):121–6. doi:10.1007/s00223-004-0122-0
2. Alifakioti, E., Arhakis, A., Oikonomidis, S. *et al.* Structural and chemical enamel characteristics of hypomineralised second primary molars. Eur Arch Paediatr Dent 22, 361–366 (2021).
3. Almuallem Z, Busuttil-Naudi A. Molar incisor hypomineralisation (MIH) - an overview. Br Dent J. Published online October 5, 2018. doi:10.1038/sj.bdj.2018.814
4. Ammar N, Fresen KF, Schwendicke F, Kühnisch J. Epidemiological trends in enamel hypomineralisation and molar-incisor hypomineralisation: a systematic review and meta-analysis. Clin Oral Investig.

2025;29(6):327. Published 2025 Jun 2. doi:10.1007/s00784-025-06411-4

5. Arnaud P, Kalabay L. Alpha2-HS glycoprotein: a protein in search of a function. Diabetes Metab Res Rev. 2002;18(4):311–4. doi:10.1002/dmrr.315
6. Bartlett JD, Smith CE. Modulation of cell-cell junctional complexes by matrix metalloproteinases. J Dent Res. 2013;92(1):10–7. doi:10.1177/0022034512463397
7. Bekes K, Mitulović G, Meißner N, Resch U, Gruber R. Saliva proteomic patterns in patients with molar incisor hypomineralization. Sci Rep. 2020;10(1):7560.
8. Elfrink ME, Ghanim A, Manton DJ, Weerheijm KL. Standardised studies on molar incisor hypomineralisation (MIH) and hypomineralised second primary molars (HSPM): a need. Eur Arch Paediatr Dent. 2015;16(3):247–55. doi:10.1007/s40368-015-0179-7
9. Elfrink ME, Veerkamp JS, Kalsbeek H, Manton DJ. Clinical characteristics and differential diagnosis of hypomineralised second primary molars and molar incisor hypomineralisation. Monogr Oral Sci. 2024;32:35–42. doi:10.1159/000538853
10. Garot E, Denis A, Delbos Y, Manton DJ, Silva M, Rouas P. Are hypomineralised lesions on second primary molars (HSPM) a predictive sign of molar incisor hypomineralisation (MIH)? A systematic review and a meta-analysis. J Dent. 2018;72:8–13
11. Garot E, Rouas P, Somani C, Taylor GD, Wong F, Lygidakis NA. An update of the aetiological factors involved in molar incisor hypomineralisation (MIH): a systematic review and meta-analysis. Eur Arch Paediatr Dent. 2022;23(1):23–38. doi:10.1007/s40368-021-00646-x.
12. Horsophonphong S, Roytrakul S, Lertruangpanya K, Kitkumthorn N, Surarit R. Proteomic analysis of dental pulp from deciduous teeth in comparison to permanent teeth: an in-vitro study. Eur Arch Paediatr Dent. 2025;26(5):941–951. doi:10.1007/s40368-025-01043-4
13. Hubbard MJ, Mangum JE, Perez VA, Nervo GJ, Hall RK. Molar hypomineralisation: a call to arms for enamel researchers. Front Physiol. 2017;8:546. doi:10.3389/fphys.2017.00546
14. K.Jähn-Rickert, J.Krug, G.Lotze, et al. “Structural Porosity and Low Mineral Density in Enamel Rods Drive Molar Incisor Hypomineralisation.” Adv. Funct. Mater. (2026): e24830. https://doi.org/10.1002/adfm.202524830
15. Lygidakis NA, Garot E, Somani C et al (2022) Best clinical practice guidance for clinicians dealing with children presenting with molar-incisor-hypomineralisation (MIH): an updated European academy of paediatric dentistry policy document. Eur Arch Paediatr Dent 23:3–21. https://doi.org/10.1007/S40368-021-00668-5
16. McCarra C, Olegário IC, O’Connell AC, Leith R. Prevalence of hypomineralised second primary molars (HSPM): A systematic review and meta-analysis. Int J Paediatr Dent. 2022;32(3):367–82.
17. Mangum JE, Crombie FA, Kilpatrick N, Manton DJ, Hubbard MJ. Surface integrity governs the proteome of hypomineralized enamel. J Dent Res. 2010;89(10):1160–5. doi:10.1177/0022034510375821
18. Mukhtar U, Goyal A, Luthra-Guptasarma M, Gauba K, Kapur A, Thakur AK. Label-free quantitative proteomics reveals molecular correlates of altered biomechanical properties in molar incisor hypomineralization (MIH): an in vitro study. Eur Arch Paediatr Dent. 2022;23(1):179–91.

19. Negre-Barber A, Montiel-Company JM, Boronat-Catalá M, Catalá-Pizarro M, Almerich-Silla JM. Hypomineralised second primary molars as predictor of molar incisor hypomineralization. J Clin Med. 2023;12(18):5875. doi:10.3390/jcm12185875
20. Pappa E, Vastardis H, Makridakis M, Zoidakis J, Vougas K, Stamatakis G, et al. Analysis of human and microbial salivary proteomes in children offers insights on the molecular pathogenesis of molar-incisor hypomineralization. Biomedicines. 2022;10(9):2061. doi:10.3390/biomedicines10092061
21. Perez VA, Mangum JE, Hubbard MJ. Pathogenesis of molar hypomineralisation: aged albumin demarcates chalky regions of hypomineralised enamel. Front Physiol. 2020;11:579015. doi:10.3389/fphys.2020.579015
22. Rexhaj F, Sabel N, Robertson A, Lundgren T. Proteomic profiling of human dental enamel affected by molar incisor hypomineralisation of different clinical severity grades: an in vitro study. Eur Arch Paediatr Dent. 2024;25(4):533–45. doi:10.1007/s40368-024-00911-9
23. Simmer JP, Hu JC. Expression, structure, and function of enamel proteinases. Connect Tissue Res. 2002;43(2–3):441–9. doi:10.1080/03008200290001159
24. Singh A, Goyal A, Luthra-Guptasarma M, Kapur A, Jaiswal M, Madan S. A proteomics approach to understand the hypomineralisation of second primary molars (HSPM). In-vitro study. Eur Arch Paediatr Dent. 2025;26(6):1085-1094.0020
25. Singh A, Tamchos R, Luthra-Guptasarma M, Goyal A, Jaiswal M, Zalmi MKK. Comparative Proteomic Profiling of Enamel of Neonatal Teeth and Normal Primary Teeth Using Liquid Chromatography-Mass Spectrometry Approach. Pediatr Dent. 2025;47(6):416–428.
26. Tang D, Chen M, Huang X, Zhang G, Zeng L, Zhang G, Wu S, Wang Y. SRplot: A free online platform for data visualization and graphing. PLoS One. 2023;18(11):e0294236.
27. Weerheijm KL, Jälevik B, Alaluusua S. Molar-incisor hypomineralisation. Caries Res. 2001;35(5):390–1. doi:10.1159/000047479

Figures

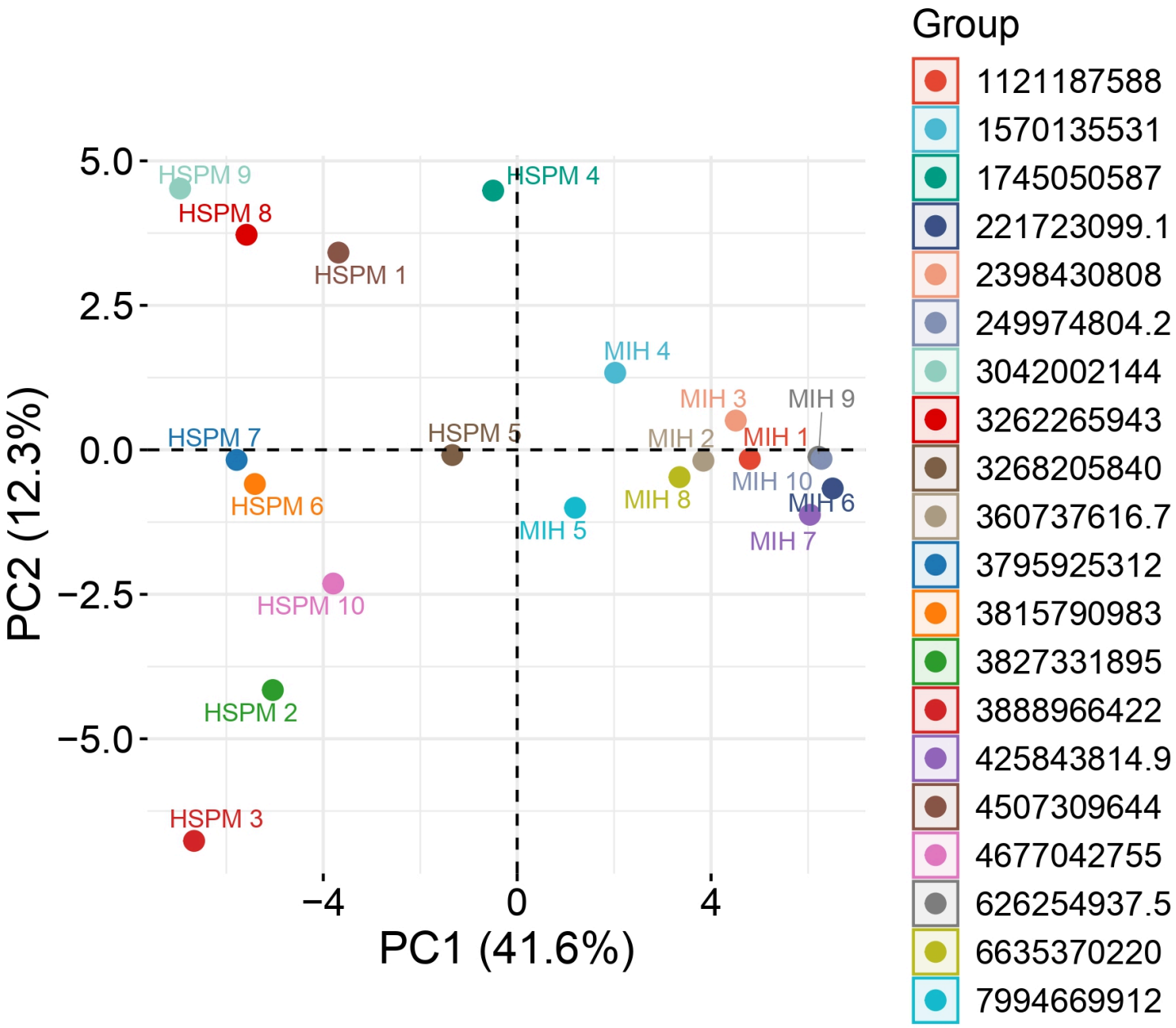


Figure 1

Fig. 1. Principal Component Analysis (PCA) of log_2-transformed enamel protein abundances.

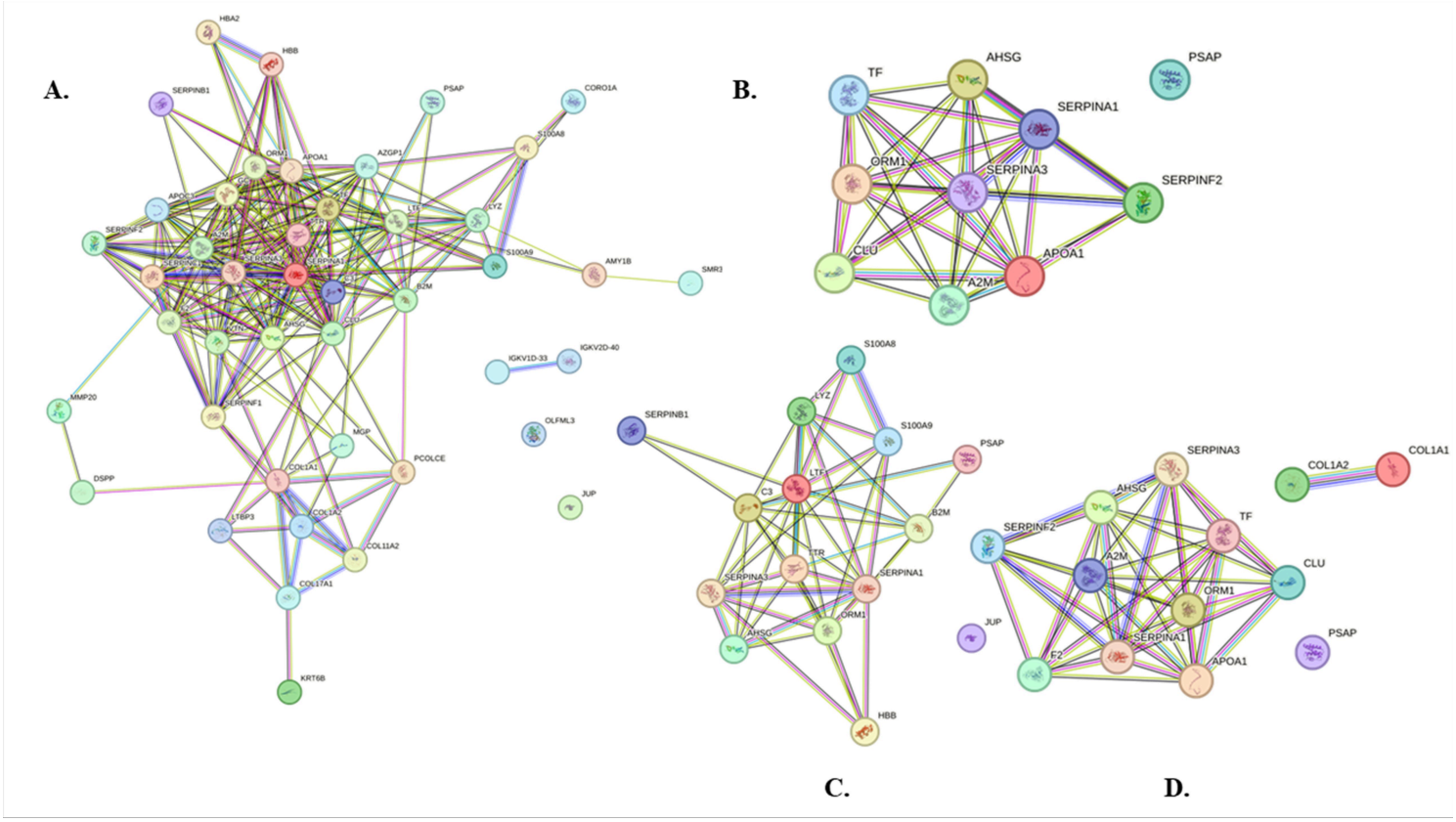


Figure 2

Fig. 2A. STRING protein–protein interaction network and REACTOME pathway enrichment analysis in HSPM affected enamel (A) showing interactions belonging to categories involved in platelet degranulation (Panel B), neutrophil degranulation(Panel C), & Platelet activation, signalling & aggregation (Panel D).

Fig. 2B. STRING protein–protein interaction network and various Gene Ontology (GO) enrichment analysis in MIH affected enamel (Panel A) showing interactions belonging to categories involved in Molecular function (Panel B), Cellular componenet (Panel C), Local Network Cluster (Panel D), KEGG pathways (Panel E) & REACTOME pathway analysis (Panel F). Further key interactions in REACTOME pathway analysis show enrichment for categories involved in cornified envelope formation (i), apoptosis (ii), oxidative stress–induced senescence (iii), and innate immune system pathways (iv).

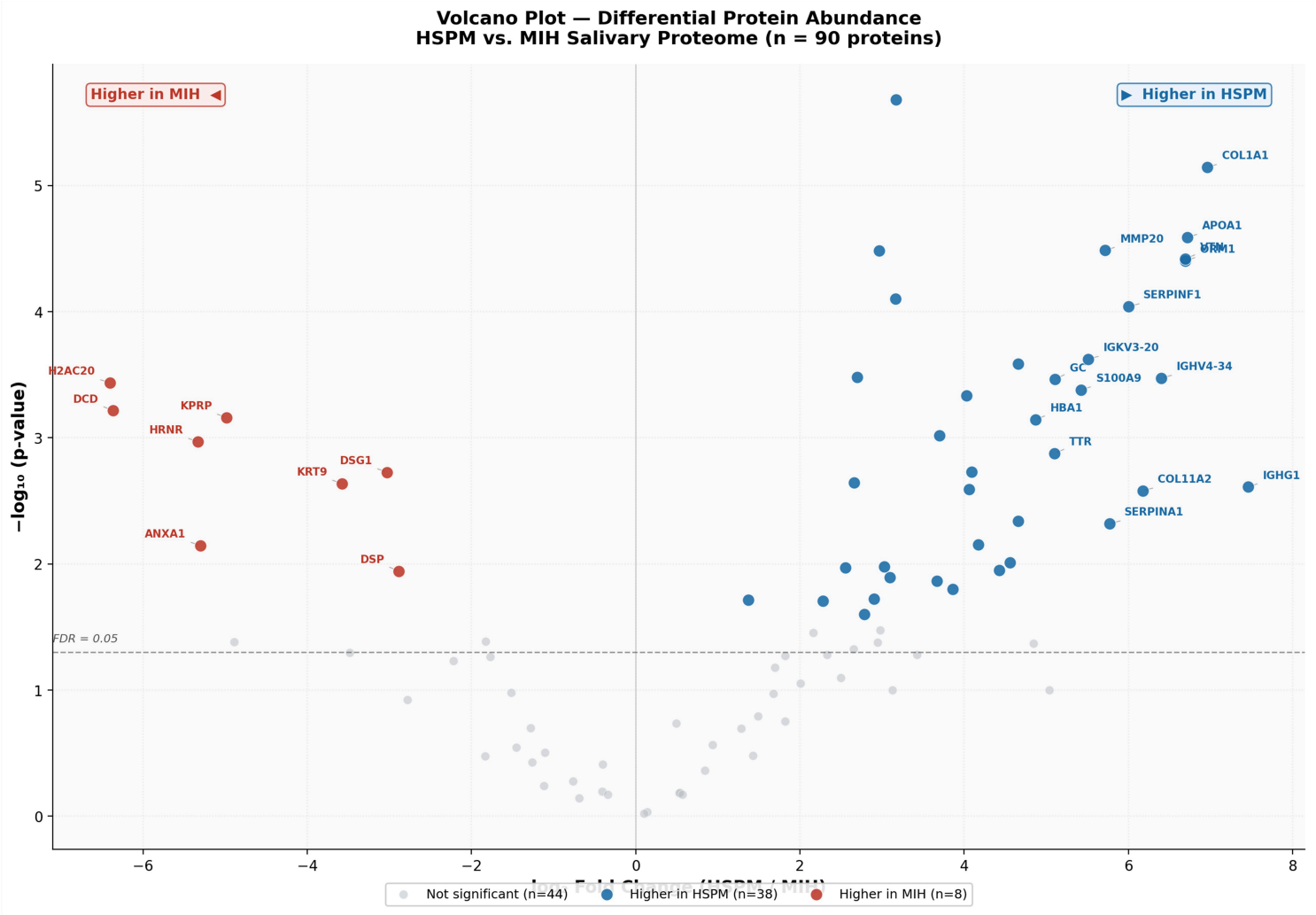


Figure 3

Fig. 3A. Volcano plot of differential enamel matrix protein abundance between HSPM and MIH

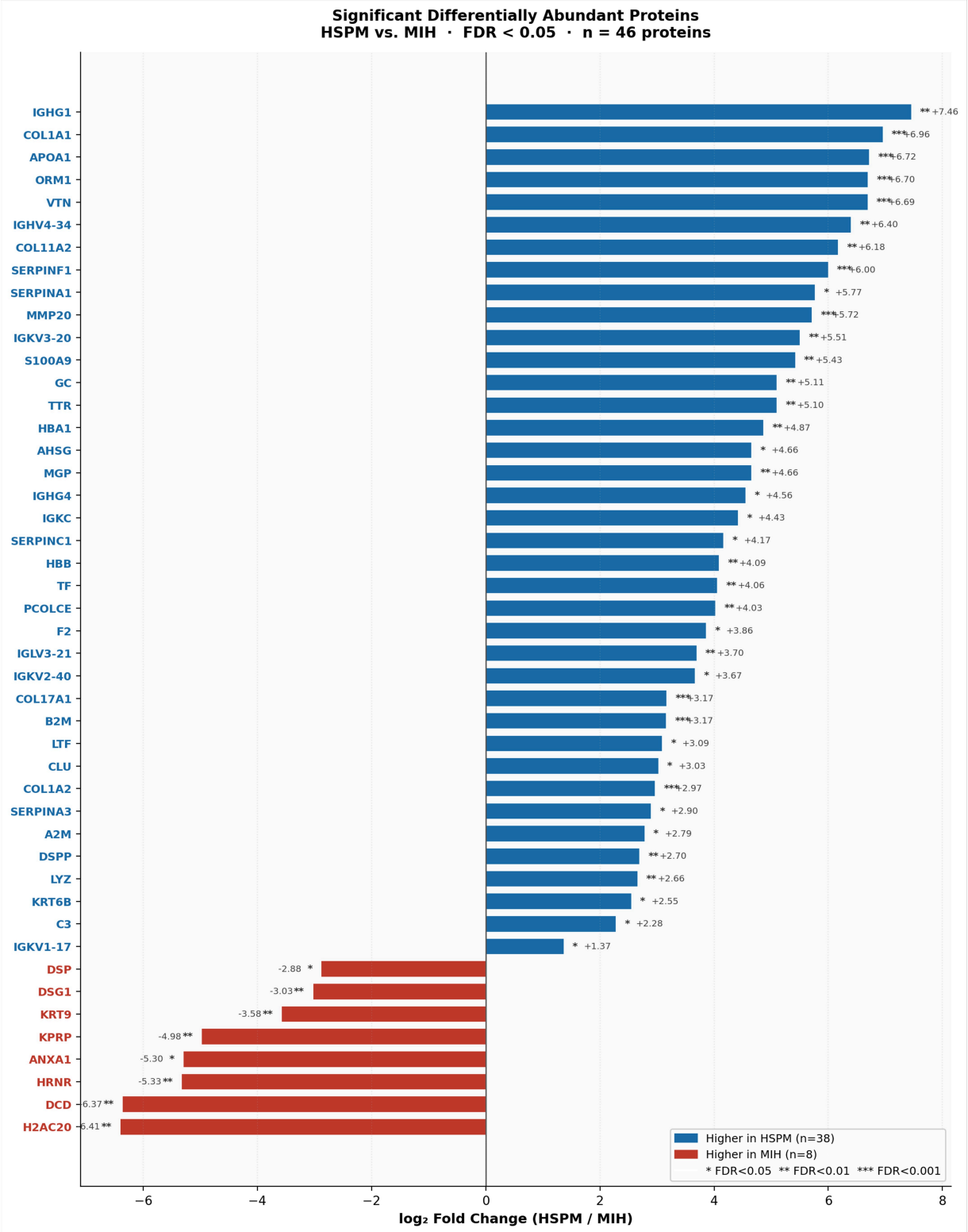


Figure 4

Fig. 3B. Ranked log_2 fold change bar chart of all significantly differentially abundant enamel proteins.

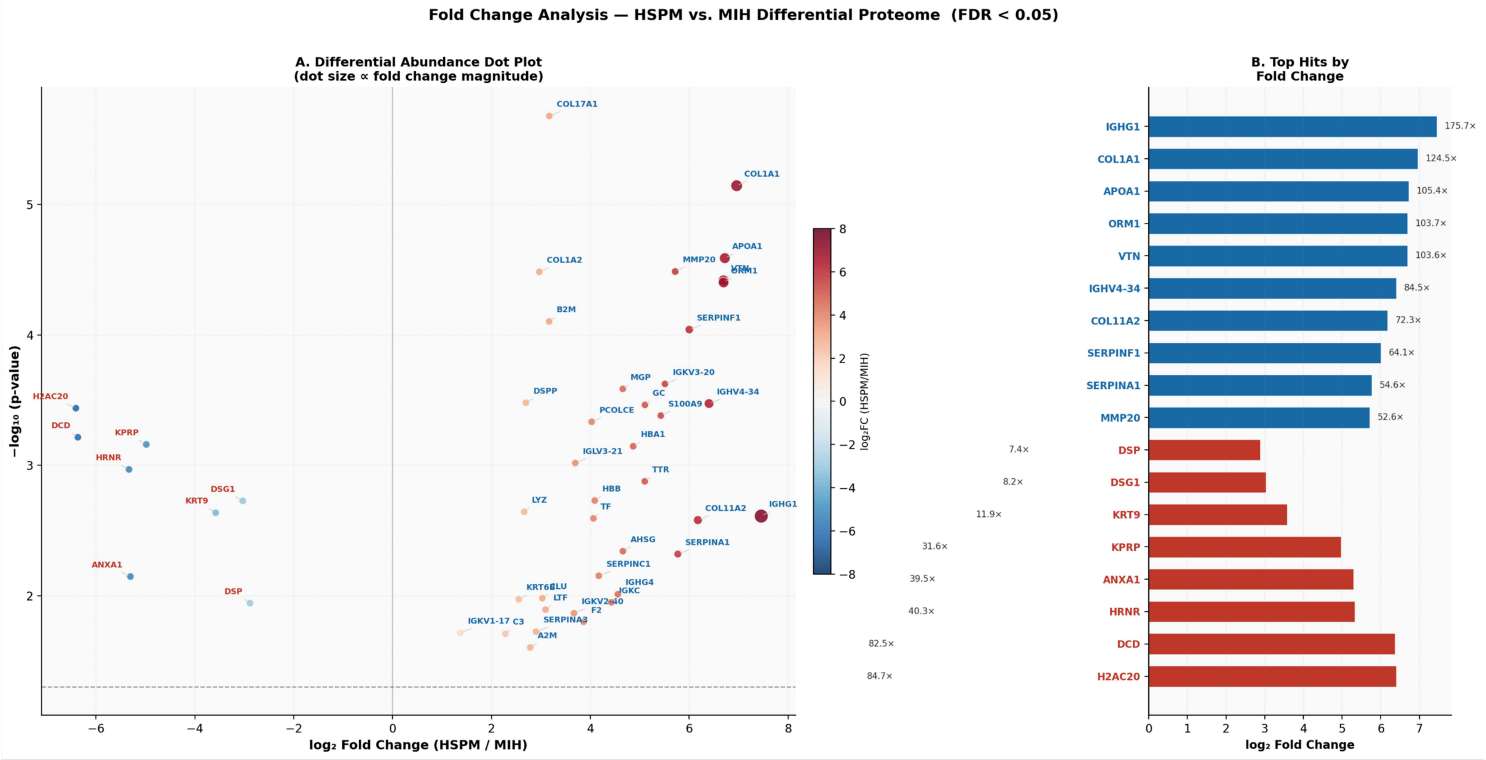


Figure 5

Fig. 3C. Differential abundance dot plot and top-hit fold change summary for HSPM vs. MIH enamel proteins.

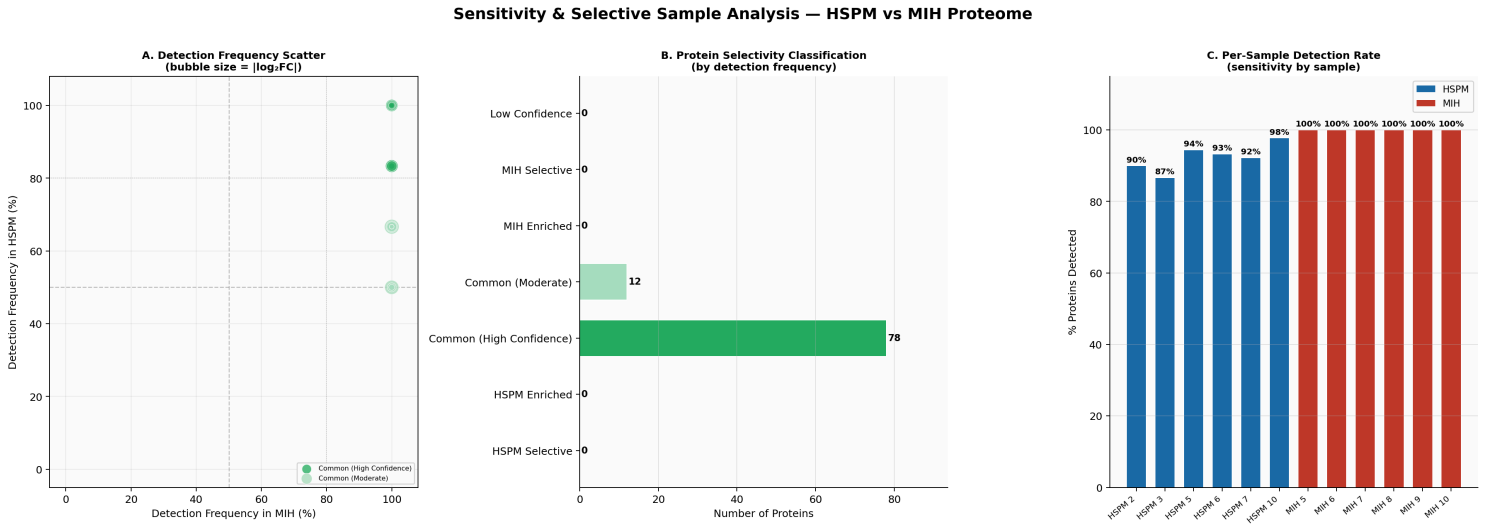


Figure 6

Fig. 4. Sensitivity analysis of the HSPM and MIH enamel proteomes : a) Detection frequency scatter b) Selectivity classification c) Per sample detection rates

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.docx
- SuplementaryFig1.png
- SupplementaryFig2.tif
- SupplementaryFigure3.tif