bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Proteomic pipeline to identify age-related compositional and structural changes of the human dentin extracellular matrix

Mariana Reis[a,b], Fred Lee[c], Ana K. Bedran-Russo[a,b*], Alexandra Naba[c*]

[a] Department of Restorative Dentistry, College of Dentistry, University of Illinois at Chicago, Chicago, IL, USA

[b] Department of General Dental Sciences, Marquette University, Milwaukee, WI, USA

[c] Department of Physiology & Biophysics, College of Medicine, University of Illinois at Chicago, Chicago, IL, USA

***Corresponding Authors*:**

* AKBR: ana.bedran-russo@marquette.edu

* AN: anaba@uic.edu

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

Objective: Devise a pipeline to investigate the protein composition of the human root dentin extracellular matrix (ECM) from single individuals of different age cohorts.

Design: Individual cervical root dentin of sound human molars from two age brackets, young (18-25 years old; n=3) and old (75-85 years old; n=3), were cut and pulverized. Protein extraction and fractionation were completed by sequential demineralization with EDTA buffer, chaotropic extraction with guanidine hydrochloride, and urea. The resulting protein extracts of differential solubility were digested into peptides and peptides were analyzed by mass spectrometry. Data generated for this study are available via ProteomeXchange, identifier PXD018320.

Results: We found that protein extracts of different solubilities present distinct biochemical compositions. We further define the matrisome of young (48 proteins) and old (50 proteins) human root dentin and report the identification of compositional and structural differences in ECM proteins from young and old teeth.

Conclusion: Our study provides a rigorous pipeline, from sample preparation to data analysis, to investigate the ECM composition – or matrisome – of the dentin. This pipeline has the potential to lead to the discovery of biomarkers of tooth aging and health.

***Keywords*:**

Matrisome, Tooth, Aging, Collagens, Proteoglycans, Mass spectrometry

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

1. Introduction

The world population is aging at a high rate. Studies predict that, by 2050, the number of adults over the age of 65 will increase by 16%, and the number of 80-year-olds and older individuals is estimated to triple (United Nations, Department of Economic and Social Affairs, 2019). These cohorts will sustain their natural dentition and experience higher prevalence of oral health conditions (Friedman et al., 2014). Particularly, root caries, a multifactorial disease on the root surface of the tooth compromises oral health and function in elderly individuals, with high prevalence in adults of 65 years old and above (Hayes et al., 2017; Zhang et al., 2020). This disease results in the progressive breakdown of the tissue as a consequence of mineral loss and degradation of the dentin extracellular matrix (ECM).

Mineralized tissues, like bone and teeth, are well-organized hierarchical structures consisting of an ECM that provides structural functionality for the deposition of an inorganic phase made of hydroxyapatite crystals (Alford et al., 2015; Marshall, 1993). With aging, the dentin tissue exhibits higher mineral content (Kinney et al., 2005), a gradual reduction of dentin tubule diameter and dental pulp size (Iezzi et al., 2019), and decreased mechanical properties (Arola & Reprogel, 2005; Ryou et al., 2015). However, there is a significant gap in our knowledge of the modifications in the composition and abundance of dentin ECM components with aging. Thus, having a cartography of the dentin ECM would be a step towards the discovery of potential biomarkers of the aging tooth. It would also be a step towards the development of strategies to maintain or regenerate a functional dentition.

Over the past decade, mass-spectrometry-based proteomic approaches have been used to study many aspects of oral biology including the profiling of the saliva and enamel pellicle (Lau et al., 2021; Rosa et al., 2012; Ventura et al., 2017) but also the composition of different parts of the tooth such as the enamel (Green et al., 2019; Hu et al., 2005; Jágr et al., 2019), the dentin (Horsophonphong et al., 2020; Jágr et al., 2012; Park et al., 2009; Widbiller et al., 2019), the pulp (Abbey et al., 2018; Eckhard et al., 2015; Jágr et al., 2016; Pääkkönen et al., 2005; Wei et al., 2008), and the cementum (Salmon et al., 2013). These studies employed diverse methodologies of sample preparation, protein extraction, identification, and quantification. For example, protein extraction was performed using different reagents, such as guanidine hydrochloride (Abbey et al., 2018; Eckhard et al., 2015; Jágr et al., 2012), radioimmunoprecipitation assay buffer (RIPA) (Park et al., 2009), ethylenediamine tetraacetic acid (EDTA), among others (Chun et al., 2011;

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Horsophonphong et al., 2020; Jágr et al., 2012; Sharma et al., 2020; Widbiller et al., 2019; Wojtas et al., 2020). The analysis of proteomic datasets has been explored using various bioinformatic tools that are critical to the interpretation of the MS/MS spectra and protein identification (Calderón-González et al., 2016; Cho, 2007).

However, while these studies identified a variety of proteins in human teeth, in many instances the level of confidence with which ECM proteins were identified lacked the rigor expected from such studies, and they were not designed to investigate specifically ECM proteins that present unique biochemical features. Moreover, among the fifteen studies assessed, and that focused on tooth proteomics, only five used human dentin as their substrate (Horsophonphong et al., 2020; Jágr et al., 2012, 2019; Park et al., 2009; Widbiller et al., 2019), and just two studies targeted the dentin from the root of the tooth (Jágr et al., 2012; Park et al., 2009). Interestingly, only one study investigated separately the protein composition of young individuals (n=3) (Park et al., 2009), while the majority of the studies pooled their samples, regardless of age, gender, or area of the tooth. Furthermore, the lack of use of a unified annotation of ECM components prevented a comprehensive definition of the dentin ECM proteins and comparisons across different cohorts and studies.

The large size and high insolubility of ECM proteins have made their biochemical analysis notoriously difficult (Taha & Naba, 2019). In the past years, we and others have developed experimental proteomic pipelines tailored to enrich, solubilize, and digest ECM proteins (Naba et al., 2012; Randles & Lennon, 2015; Schiller et al., 2015; Taha & Naba, 2019). We have also used bioinformatics to define the "matrisome", the collection of genes encoding ECM and ECM-associated proteins, and have shown that the matrisome list can be used to comprehensively annotate high-throughput data for ECM components (Gebauer & Naba, 2020; Naba et al., 2016; Naba & Ricard-Blum, 2020). With these tools in place, proteomics has now become a method of choice to study the changes in the ECM associated with diseases such as cancers (Socovich & Naba, 2019), fibrosis (Schiller et al., 2019) or during aging (Angelidis et al., 2019; Grilo et al., 2020).

Here, we assessed the feasibility of using proteomics to characterize the ECM composition of the root dentin from single individuals. To do so, we devised an experimental and stringent analytical pipeline to compare the ECM proteins found in the soluble and insoluble components of the human cervical root dentin. Using this pipeline, we first assessed the extent of inter-

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

individual variability in cohorts of same age and gender. We then compared the ECM protein composition of root dentin from young and older adult individuals. Last, we showed that we could gain structural insight from the proteomic data. Altogether, our study provides a pipeline to investigate the composition of the dentin ECM that has the potential to lead to the discovery of biomarkers of the aging tooth.

2. Material and methods

2.1. Sample selection and preparation

Three sound molars were selected from two age brackets, young (18-25 years old) and older (75-85 years old) men, according to protocol 2018-0346 approved by the Institutional Review Board Committee of the University of Illinois at Chicago (**Supplementary Table 1A**). Teeth were immediately stored after extraction at -20ºC. The cervical part of the root (**Fig. 1A**) was obtained by sectioning the tooth with an IsoMet™ high precision cutting machine (Buehler Ltd, Lake Bluff, IL, USA). All soft tissues, periodontal ligament, cementum, and pulp tissue were carefully removed with curettes and endodontic files.

2.2. Dentin protein extraction

The extraction of dentin proteins was adapted from an existing protocol (Martin-De Las Heras et al., 2000; Park et al., 2009), with modifications (**Fig. 1B**). In brief, the cervical dentin root sections (250-350 mg per tooth/sample) were pulverized, and the powder demineralized with a 0.5M EDTA solution at pH 8.0 containing protease inhibitors (Pierce Protease Inhibitor Tablets, EDTA-free, Thermo Fisher Scientific, Rockford, IL, USA) for 12 days under agitation at 4°C. The EDTA solution was changed every 3 days. Solubilized proteins were obtained by centrifugation at 5000 x g for 10 min at 4°C. The resulting supernatant containing EDTA-soluble proteins will be subsequently referred to as the “E-extract”. Protein extraction from the remaining EDTA-insoluble pellet was performed with a 4M guanidine hydrochloride (Gu-HCl, Sigma-Aldrich St. Louis, MO, USA) solution at pH 7.4 containing protease inhibitors, for 3 days under agitation at 4°C. Solubilized proteins were obtained by centrifugation at 5000 x g for 10 min at 4°C. The resulting supernatant containing Gu-HCl-soluble proteins will be subsequently referred to as the “G-extract”. Proteins from the remaining Gu-HCl-insoluble pellet will be further termed the dentin

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

extract or "D-extract". The E- and G-extracts were dialyzed against ultrapure water using M_r 6,000 cut-off Spectra/Por dialyses tubing (Spectrum Laboratories, Rancho Dominguez, CA, USA) at 4°C for 6 h. The ultrapure water was exchanged twice. The total concentration of proteins was measured with the Pierce Rapid Gold BCA Protein assay kit (Thermo Fisher Scientific, Rockford, IL, USA). The E-, G-, and D-extracts were lyophilized (FreeZone Freeze Dryer, Labconco, Kansas City, MO, USA) and stored at -20°C for further analyses.

2.3. Protein digestion

The lyophilized samples (E-extracts, G-extracts and D-extracts, 5-10 mg dry weight) were subsequently solubilized and digested into peptides following an established and previously described protocol (**Fig. 1B**) (Naba et al., 2012). Briefly, the proteins were solubilized in 8M urea, and protein disulfide bonds were reduced using 10mM dithiothreitol and alkylated with 25mM iodoacetamide (Thermo Fisher Scientific, Rockford, IL, USA). Proteins were then deglycosylated with PNGaseF (New England Biolabs, Ipswich, MA, USA) and digested with Lys-C (Thermo Fisher Scientific, Rockford, IL, USA) for 2 hours at 37°C, and trypsin (Thermo Fisher Scientific, Rockford, IL, USA), overnight at 37°C. The samples were acidified with 50% trifluoroacetic acid (TFA, Thermo Fisher Scientific, Rockford, IL, USA) until the pH of the samples was lesser than 2. Acidified samples were desalted in C18 desalting columns (Pierce Peptide Desalting Spin Columns, Thermo Fisher Scientific, Rockford, IL, USA) and reconstituted in 95% HPLC grade water, 5% acetonitrile and 0.1% formic acid. Peptide concentration was measured with a colorimetric assay (Pierce Quantitative Colorimetric Peptide Assay kit, Thermo Fisher Scientific, Rockford, IL, USA).

2.4 Analysis of digested dentin proteins by LC-MS/MS

Approximately 300 ng of desalted peptides were analyzed at the University of Illinois Mass Spectrometry core facility on a Thermo Fisher Orbitrap Velos Pro coupled with Agilent NanoLC system (Agilent, Santa Clara, CA). The LC columns (15 cm × 75 µm ID, Zorbax 300SB-C18) were purchased from Agilent. Samples were analyzed with a 120-min linear gradient (0–35% acetonitrile with 0.1% formic acid) and data were acquired in a data-dependent manner, in which MS/MS fragmentation was performed on the top 10 intense peaks of every full MS scan. Full MS scans were acquired in the Orbitrap mass analyzer over m/z 350–1800 range with resolution 30,000

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(m/z 400). The target value was 1.00E+06. The ten most intense peaks with charge state ≥ 2 were fragmented in the HCD collision cell with normalized collision energy of 35%, these peaks were then excluded for 30 s after 2 counts within a mass window of 10 ppm. Tandem mass spectrum was acquired in the Orbitrap mass analyzer with a resolution of 7,500. The target value was 5.00E+04. The ion selection threshold was 5,000 counts, and the maximum allowed ion accumulation times were 500 ms for full scans and 250 ms for HCD.

All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.6.2). Mascot was set up to search the Uniprot-human_20190917 database (20,430 entries; individuals 1 and 6) Uniprot-human_20191015 database (20,365 entries; individuals 2, 3, 4, 5) assuming stricttrypsin. Mascot was searched with a fragment ion mass tolerance of 0.30 Da and a parent ion tolerance of 15 PPM. Carbamidomethyl of cysteine was specified in Mascot as a fixed modification. Gln->pyro-Glu of the N-terminus, deamidated of asparagine and glutamine and oxidation (hydroxylation) of methionine, proline, and lysine were specified in Mascot as variable modifications.

Scaffold (version Scaffold_5.0.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. The following stringent thresholds were applied: peptide identifications were accepted if they could be established with a false-discovery rate (FDR) of less than 1.0% by the Peptide Prophet algorithm (Keller et al., 2002) with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least 2 identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al., 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters.

Mass spectrometry output was further annotated to identify ECM and non-ECM components using the human matrisome list (Naba et al., 2012, 2016). In brief, matrisome components are classified as core-matrisome or matrisome-associated components, and further categorized into groups based on structural and/or functional features: ECM glycoproteins, collagens or proteoglycans for core matrisome components; and ECM-affiliated proteins, ECM regulators, or secreted factors for matrisome-associated components (Hynes & Naba, 2012; Naba et al., 2012).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Raw mass spectrometry data have been deposited to the ProteomeXchange Consortium (Deutsch et al., 2020) via the PRIDE partner repository (Perez-Riverol et al., 2019) with the dataset identifier PXD018320 and 10.6019/PXD018320. *The raw data will be made publicly available upon acceptance of the manuscript*.

3. Results

3.1. Feasibility of the analysis of samples from single individuals

To assess the feasibility of using proteomics to characterize the ECM composition of root dentin from each individual, we used the cervical area of the root of sound molars from young individuals (18-25 years old) and fractionated it to generate protein samples of different solubility (**Fig. 1B**). The sample preparation protocol yielded, for each sample, three distinct fractions, an EDTA-soluble-extract "E-extract", a Gu-HCl-soluble-extract "G-extract", and a remaining pool of insoluble dentin proteins, enriched for ECM proteins, further termed the "D-extract" (**Fig. 1B**). A total of 5-10 mg (in dry weight) of each protein extract was solubilized and digested into peptides as previously described (Naba et al., 2012). After digestion and desalting, the amount of peptide obtained for each type of extracts ranged from of 10-140 µg for the E-extracts, 13-189 µg for the G-extracts and 155-343 µg for the D-extracts. We observed that the least amount of peptides were obtained from the most soluble E-extract while the most insoluble protein fraction, D-extract, contained the most peptides. Nevertheless, sufficient amount of peptides were obtained from each extract for mass spectrometric analysis. This demonstrates the feasibility of our pipeline, where a sample from a single tooth and from a single individual provides enough material for the downstream proteomic analysis.

While many of the dentin proteomic studies focused on the analysis of the most soluble protein fractions, the goal of this study was to evaluate both soluble and insoluble components, and thus assess the inter-individual variability in the protein composition and solubility of the human cervical root dentin within cohorts of same age and gender. To streamline the reporting of our results, we further combined the identified protein fractions of higher solubilities, the E- and G- extracts, into one larger group subsequently termed "supernatant" and compared these lists to the ECM proteins identified in the insoluble dentin pellet (*i.e.*, resistant to EDTA and Gu-HCl solubilization), the "D-extract".

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

3.2. Protein composition of extracts of different solubilities

By considering ECM proteins detected with at least two peptides in at least one sample, a total of 99 proteins were detected in the cervical area of root dentin of young individuals (**Supplementary Table 1B**), 48 of which were matrisome proteins (**Supplementary Tables 1C and 1D**). Initially, we compared the matrisome proteins of supernatants and D-extracts from young dentin samples (n = 3). We found 38, 42 and 39 ECM proteins from individuals 1, 2, and 3, respectively (**Fig. 2A and Supplementary Tables 1D-G**). Collectively, supernatants (33-38 proteins) presented a higher number of identified ECM proteins than D-extracts (24-27 proteins) The number of ECM proteins detected in both supernatants and D-extracts was similar among the three individuals, an average of 21 proteins (**Supplementary Tables 1D-G**). We also observed that proteins fractions of higher solubility, *i.e.*, the supernatants, contained a higher number of uniquely identified ECM proteins, 12 to 17 (36-45%), in comparison to the D-extracts (2 to 6 proteins or 8-22% of the proteins detected) (**Fig. 2A**). The analysis of the insoluble D-extracts revealed that 50% of the identified ECM proteins are collagens and that half as many glycoproteins, proteoglycans, and secreted proteins were detected compared to the supernatants (**Supplementary Tables 1D-G**).

3.3. Assessment of inter-individual variability

We next sought to assess the extent of inter-individual variability of dentin proteins between three distinct young individuals. We collated the proteins identified in either supernatants or D-extracts for each individual and compared these lists (**Fig. 2B-C and Supplementary Tables 1D-G**). The union of three lists revealed 46 and 32 ECM proteins were identified respectively in the supernatants and D-extracts of young individuals (**Fig. 2B-C and Supplementary Table 1H**). We identified a set of 25 ECM proteins in supernatants that intersect across all individuals (**Fig. 2B**), and 18 common ECM proteins in D-extracts (**Fig. 2C**). The number of uniquely identified ECM proteins found in each young individual varied according to the fraction, where a higher number of proteins was detected in supernatants (10), compared to D-extracts (6), indicating a greater variability in samples of higher solubility.

3.4. Defining the young dentin matrisome

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

To further characterize the ECM composition of the young cervical dentin, we defined the dentin "matrisome" as the ensemble of proteins detected in at least one individual and in either supernatants or D-extracts. Using this definition, the dentin matrisome of young adult individuals comprises of 48 ECM proteins (**Fig. 2D, Table 1, Supplementary Table 1H**). The dentin ECM composition of the young cohort consisted of predominantly 18 collagens (38%) and 13 glycoproteins (27%), along with 5 proteoglycans (10%), where, together, they constitute the core matrisome. Matrisome-associated proteins were also present, however in a smaller quantity, represented by 1 ECM-affiliated protein (2%), 8 ECM regulators (17%), and 3 secreted factors (6%) (**Fig. 2D**).

The main proteins that have previously been shown to compose the dentin ECM were detected. Collagens, as expected, were the most abundant protein in the young dentin matrisome. The two chains forming the fibrillar type I collagen (COL1A1 and COL1A2) were predominant and the most abundant protein detected in our sample (**Supplementary Table 1D, columns O – W** representing normalized total spectral count, a proxy for protein abundance), however, we also report the identification of a number of collagens not previously associated with the dentin tissue, including collagens type IV-V, VII, IX, XVI, XXII, XXV, and XXVIII (**Table 1 and Supplementary Table 1H**).

We identified glycoproteins such as dentin sialophosphoprotein (DSPP), osteopontin (SPP1), procollagen C-endopeptidase enhancer (PCOLCE), secreted protein acidic and cysteine rich (SPARC), matrix Gla protein (MGP) and vitronectin (VTN) that are known dentin proteins with important roles in biomineralization, enzymatic cleavage of type I procollagen and cell adhesion (Kawasaki et al., 2004; Linde, 1989; Ravindran & George, 2014; Ritchie, 2018). We also report the detection of other glycoproteins never found before in the root dentin ECM, such as collagen triple helix repeat-containing protein 1 (CTHRC1) and fibrillin 1 (FBN1), the later detected in both supernatant and D-extracts, and the last exclusively in D-extracts.

Small leucine-rich proteoglycan (SLRP) family were the major proteoglycans (PGs) found. They include biglycan (BGN) and asporin (ASPN) (Class I), lumican (LUM) and osteomodulin (OMD) (Class II), and chondroadherin precursor (CHAD) (Class IV), which are known to participate in collagen fibril assembly, tissue growth and regulation of interfibrillar spacings (Iozzo & Schaefer, 2015).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

ECM-associated proteins were identified including the coagulation factors (F2, F9 and F10), and a matrix metalloproteinase protein (MMP20) that are involved in the breakdown of extracellular matrix in normal physiological processes, tissue remodeling/degradation, and disease processes (Loffek et al., 2011). Insulin like growth factor 2 (IGF2), a member of the insulin family of polypeptide growth factors and involved in development and growth of tissues (Baxter, 2000) was identified for the first time in the root dentin.

We further aimed to compare the list of ECM proteins identified here with those reported in previous investigations that explored the proteome of human root dentin (Jágr et al., 2012; Park et al., 2009). To compare our data with the earlier studies, we retrieved and reannotated them with the *in silico*-predicted complete matrisome list (Naba et al., 2016). The data are presented in **Supplementary Table 2A**. The assessment revealed 29 ECM proteins found in all three studies and another set of 19 ECM proteins that were detected in two out of the three studies (**Fig. 2E and Supplementary Table 2B**). The presence of proteins such as dermatopontin (DPT), dentin sialophosphoprotein (DSPP), biglycan (BGN), and matrix metallopeptidase-20 (MMP20) further confirms the established association of these proteins to the dentin. However, a greater difference is revealed between the three matrisomes: while 35 and 25 proteins were unique to the two studies (Jágr et al., 2012; Park et al., 2009), respectively, our study identified an additional 13 ECM proteins not detected before in the human root dentin. These 13 structural ECM proteins consisted of 10 collagens, 2 glycoproteins and 1 secreted factor similarly present in young supernatants and D-extracts (**Fig. 2E and Supplementary Table 2**). This meta-analysis demonstrates the strength of our ECM-focused proteomic and analytical pipeline in the cervical area of root dentin and reveals an additional list of collagens never identified in this type of sample.

Since our approach allowed for the identification of proteins from individual dentin samples of different solubilities, we further compared the samples of similar solubilities from the 2012 study by Jágr et al. to our respective E-, G-, and D-extracts. (**Fig. S1 and Supplementary Tables 2C-E**). A set of 22 ECM proteins were identified in E-extracts of both studies (**Fig. S1A**). G-extracts and D-extracts revealed 22 and 11 ECM common proteins, respectively (**Fig. S1B-C**). While the comparison of G-extracts revealed a similar number of proteins that were uniquely identified in each study (22 – young matrisome and 33 – Jágr et al., 2012), the D-extract of the young matrisome, contained a significantly greater number of ECM proteins identified (21) when compared to Jágr et al (1), (**Fig. S1C**). The compositional differences noticed in the D-extracts are

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

remarkable and indicate that advances in the enrichment of the ECM and protein solubilization might have had a significant impact on protein identification.

Of note, the dentin tissue is part of the dentin-pulp complex, where the odontoblast layer forms the outermost layer of the pulp and is adjacent to the predentin, an initial stage of new dentin tissue. Odontoblast processes extend into the dentinal tubules (Pashley, 1996). As such, proteins from this tissue could overlap with the dentin-pulp complex tissue, and even possibly contaminate the dentin tissue. We thus sought to compare the young dentin matrisome with data from prior studies on the proteome of the pulp tissue and/or odontoblast layer (Abbey et al., 2018; Pääkkönen et al., 2005; Wei et al., 2008). The comparison with Pääkkönen study on pulp showed only one protein, SERPINA1, was also found in the young dentin matrisome defined here, while we found no overlap between our study and the Wei et al., 2008 study. A total of 153 matrisome proteins were identified as part of the pulp and 102 matrisome proteins as part of the odontoblast layer, 84 of these being found in both compartments (Abbey et al., 2018). Our comparison revealed that 27 of the 153 proteins found in the pulp were also part of the young dentin matrisome and 24 of the 102 proteins of the odontoblast layer overlapped with the young dentin matrisome (**Fig. S1D and Supplementary Table 2F**). Importantly, 18 proteins of the young dentin matrisome were neither identified in the pulp nor the odontoblast layer. These comparisons demonstrate that there are distinct ECM signatures that define the different layers of the tooth even though a substantial association is seen in the tissues of the dentin-pulp complex.

3.5. Defining the old dentin matrisome

With our pipeline in place, we next compared the ECM composition of the root dentin of young individuals (*see above*) to that of older individuals. To do so, we analyzed samples from 3 older adults (75-85 years old). A total of 101 proteins in the cervical area of root dentin of old individuals was identified (**Supplementary Table 1B**), where 50 were matrisome proteins (**Supplementary Tables 1C and 1I**). The total number of ECM proteins identified from individuals 4-6 was 43, 40 and 45, respectively (**Fig. 3A and Supplementary Tables I-L**). The number of shared ECM proteins was comparable between individuals, where an average of 16 proteins identified in supernatants and D-extracts. Conversely, the number of ECM proteins found in either fraction varied according to the fraction and individual. The supernatants of two out of three old individuals showed a higher number of unique ECM proteins, 24 - 27 (55%), in

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

comparison to 3 unique ECM proteins (7%) detected in the D-extracts (**Fig. 3A**). Individual 5 showed in opposing trend, however with a comparable number of total and overlapping ECM proteins between fractions.

The inter-individual variability across older samples was evaluated and we identified ECM proteins in either supernatants or D-extracts for each individual and compared these lists (**Fig. 3B**). Our analysis identified 50 proteins in the supernatants (**Fig. 3B**) and 43 in the D-extracts (**Fig. 3C**). A set of 17 ECM proteins was identified in supernatants that overlap between all individuals, and 12 ECM common proteins in D-extracts. The number of ECM proteins uniquely identified in older individuals varied given the fraction, where a higher number of ECM proteins was detected in D-extracts (26), compared to those in the supernatants (12), revealing an increased variability in samples of higher insolubility.

Last, we defined the old dentin matrisome as the ensemble of proteins detected in at least one individual and in either supernatants or D-extracts that comprises of a total of 50 ECM proteins (**Table 2 and Supplementary Table M**). The old matrisome is thus constituted of a majority of core matrisome proteins, 18 collagens (36%), 12 ECM glycoproteins (24%) and 5 proteoglycans (10%). Fibrillar type I collagen (COL1A1 and COL1A2) was the most abundant protein detected in our sample (**Supplementary Table 1I, columns O – W** representing normalized total spectral count, a proxy for protein abundance). The matrisome-associated proteins were less evident and comprised 2 ECM-affiliated proteins (4%), 8 regulators (16%), and 5 secreted factors (10%) (**Fig. 3D and Table 2**).

3.6. Comparing the root dentin ECM of young and older individuals

Having defined the "old root dentin matrisome", our dataset now allows us to address the following questions: 1) can we identify changes between the ECM composition of young and old root dentin; 2) is there a difference in the solubility of the ECM proteins of young and older root dentin; and 3) can proteomic data, in particular analysis of sequence coverage, be used to infer difference in ECM protein digestibility, which could be used as a proxy for evaluating structural accessibility and protein folding.

3.6.1. Comparing the ECM protein composition of young and old dentin samples

The comparison of the two age groups revealed that a large portion of ECM proteins (46) are identified in young and old samples (**Fig. 4A-B**). The major components found in both groups

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

were 17 collagens (39%) and 11 ECM glycoproteins (26%). A smaller number of 5 proteoglycans (11% of the matrisome proteins) and ECM-affiliated proteins (1) and secreted factors (3) were detected. Overall, core matrisome proteins (33) were predominant (76%) in the ECM of both age groups, when compared to matrisome-associated proteins, 16 (15%). A set of 2 and 4 ECM proteins were uniquely found in young and old cervical root dentin, respectively. **Fig. 4A** shows the ECM proteins uniquely identified in young and old dentin matrisomes. Insulin-like growth factor-binding protein 3 (IGFBP3) and lysyl oxidase (LOX), were exclusively found in the young matrisome, while hemopexin (HPX), antithrombin-III (SERPINC1), protein S100-A9 (S100A9) and transforming growth factor beta-2 proprotein (TGFB2) were uniquely found in the old matrisome. The comparison between young and old matrisomes ultimately reveals that, while sharing a large number of proteins (46), each matrisome presents certain unique components that may relate to the particular functions of the dentin tissue. We can anticipate that expanding the sample size, may lead to the identification of additional proteins, whose abundance change with age. Beyond identification, our proteomic approach can reveal semi-quantitative differences in protein abundance (see columns reporting normalized total spectral counts in **Supplementary Tables 1B, C, D and I**). While our analysis did not reveal significant variation in abundance with aging, quantitative proteomic methods, such as those relying on peptide labeling, could provide further insights.

3.6.2. Comparing the ECM protein solubility of young and old dentin samples

Our results also show that a higher number of proteins are identified in the more soluble protein fractions (supernatants), regardless of the age group, when compared to the more insoluble fractions (D-extracts). Additionally, we found that half or more ECM proteins are common to both type of samples. A large fraction of matrisome proteins in supernatants (44) are detected in both young and old cohorts (**Fig. 4C**). The main protein category identified was collagen, represented by 17 different collagens of the total proteins (38.6%). ECM glycoproteins accounted for 11 of the shared proteins (25%), where insulin like growth factor binding protein 3 (IGFBP3) and fibrillin 1 (FBN1) were uniquely identified in young and old supernatants, respectively. A total of 5 proteoglycans were identified in both young and old cohorts' supernatants and a total of 11 matrisome associated proteins (25%). Interestingly, we identified lysyl oxidase (LOX) uniquely in young samples and hemopexin (HPX), antithrombin-III (SERPINC1), protein S100-A9 (S100A9) and transforming growth factor beta-2 (TGFB2) in old supernatants (**Table 3**).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Variability of the identified ECM proteins increased with age in the more insoluble fraction (D-extract), where 29 ECM proteins were found in both age cohorts, approximately 20% less common ECM proteins when compared to common proteins in the supernatants. A higher number of unique ECM proteins was identified in old D-extracts (14) when compared to young (3) (**Fig. 4C**). There was an increase in the number of identified ECM glycoproteins, proteoglycans and ECM regulators in D-extracts with age. Collagens such as COL12A1 and COL4A2, ECM glycoproteins like dentin sialophosphoprotein (DSPP), matrix gla protein (MGP), secreted protein acidic and cysteine rich (SPARC), and THBS1, and proteoglycans asporin (ASPN) and lumican (LUM) were found exclusively in old D-extracts, while fibrillin 1 (FBN1), osteopontin (SPP1), and dermatopontin (DPT) were uniquely observed in young D-extracts (**Table 3**).

3.6.3. Proteomics provides potential insights into structural changes of dentin ECM proteins during aging

Bottom-up proteomics, which infers protein identity based on the identification of peptides, has also the potential to shed light on protein structure via peptide location fingerprinting (Eckersley et al., 2018, 2020). Indeed, structural changes of long-lived proteins, like ECM proteins, for example during tissue aging, can result in altered proteolytic susceptibility which can be assessed by evaluating protein-sequence coverage (**Supplementary Table 1C, columns AP to BG**) and mapping identifying peptides to primary protein structures. Here, we sought insight hinting at possible changes in ECM protein structures as a function of age. While our sample size limited our statistical analysis, we could identify ECM proteins whose coverage, and hence possibly structure, varied with age.

Biglycan (BGN) has previously been associated with mineralization (Haruyama et al., 2009). We detected peptides that cover on average 38% of the protein sequence in young insoluble dentin samples (D-extracts), while the sequence coverage only reached 14.5% on average across older insoluble dentin samples (**Fig. 5 and Supplementary Table 1C, row 35, and Supplementary Table 1N**). Conversely, we detected peptides covering on average 18.3% of the protein sequence in young EDTA-soluble dentin samples, while the sequence coverage only reached 27.2% on average across older EDTA-soluble dentin samples (**Supplementary Table 1C, row 35, and Supplementary Table 1N**). Moreover, we observed peptides that were either uniquely identified in young D-extracts or old D-extracts (**Fig. 5**). As the exclusive detection of these peptides in one sample type or the other could indicate greater accessibility to trypsin, we

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

propose that a study on a larger number of samples may reveal structural differences that occur with the aging process.

Another such example is the α2 chain of type V collagen (COL5A2). We detected peptides covering on average 12.8% of the protein sequence in young insoluble dentin samples, while the sequence coverage only reached 7.8% on average across older insoluble dentin samples. Conversely, we detected peptides covering on average 3.8% of the protein sequence in young EDTA-soluble dentin samples, while the sequence coverage only reached 8.4% on average across older EDTA-soluble dentin samples (**Supplementary Table 1C, row 9**).

Insulin-like growth factor binding protein-5 (IGFBP5) is known for its role in regulating cell growth, proliferation, and differentiation in different tissues, specifically in osteoblasts and odontoblasts (Aizawa et al., 2019; Mukherjee & Rotwein, 2008). We detected peptides covering on average 13.7% of the protein sequence in young EDTA-soluble dentin samples, while the sequence coverage only reached 4% on average across older EDTA-soluble dentin samples (**Supplementary Table 1C, row 26**).Results also reveal a relatively low coverage of minor ECM components of the dentin, as compared to large, structural, ECM proteins, such as the α1 and α2 chains of type I collagen whose sequence coverages reaches on average 70% across fractions. This highlights how much more remains to be there is to learn about the dentin matrisome.

4. Discussion

In this study, we demonstrate the feasibility of using proteomics to characterize the ECM composition of human cervical root dentin from single individuals and identify age-related differences in the composition and even possibly the structure of dentin ECM proteins.

In order to acquire an adequate amount of material, previous dental proteomic studies pooled samples from different samples (Abbey et al., 2018; Jágr et al., 2012, 2016, 2019; Park et al., 2009; Salmon et al., 2013). The protein analysis of small biological samples, such as tissue fragments and cells are an important challenge in the field of proteomics, however, advances in sample preparation and MS/MS techniques enabled highly effective studies and protein identification. Here, we show that the amount of material obtained from the cervical area of root dentin from one tooth, and thus from a single individual, is sufficient to conduct a rigorous proteomic analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Prior studies have proposed that analyzing pools of ECM proteins of high, intermediate, and low solubility from tissue could shed light on their degree of assembly in the ECM scaffold (most insoluble fraction) and hence on their structural and functional properties (Barrett et al., 2018; Schiller et al., 2017). Indeed, ECM protein solubility is an important factor of ECM homeostasis and biology and can further influence the mechanical properties of tissues. Multiple mechanisms contribute to determining the level of assembly (or insolubility) of ECM proteins, including post-translational modifications such as cross-linking or, on the contrary, protein cleavage or degradation (Madzharova et al., 2019; Rogers & Overall, 2013). Here, we applied such approach, for the first time to the study of a mineralized tissue and found that protein fractions of different solubility have different biochemical composition.

Our study design also enabled the assessment of the extent of inter-individual variability within cohorts of the same age group. Our hypothesis was that variability within the ECM protein composition of young root dentin is limited. Nonetheless, it is important to acknowledge that the human oral microbiome is highly singular, and can be affected by lifestyle, hygiene, environment, genetics, diet, and disease (Willis & Gabaldón, 2020). Studies of mineralized tissues such as bone have shown that the effects of fluoride intake (Robinson & Kirkham, 1990), diet and the acid/base balance can lead to changes in tissue development, i.e., alterations in the abundance of collagen in the organic matrix (Balasse et al., 1999; Seibel, 2005). Therefore, a certain variability among the ECM protein composition of different individuals was anticipated. As an effort to limit this variability, samples were carefully selected to be within stringent age ranges and of same gender. Results show limited inter-individual variability and comparable protein composition between extracts of similar biochemical properties. Of note, studies on bone, another mineralized tissue, have found differences between samples of female and male individuals and indicate that inter-individual variability could be associated to hormonal differences between samples of different gender (Almeida et al., 2017). It would thus be interesting, in the future, to expand this study to samples from female individuals.

Comparison of the data from this study with previous studies revealed a substantial similarity between human dentin proteins. Yet, we report here the identification of a set of 13 ECM proteins never yet described in the human root dentin. The differences observed in the number of identified ECM proteins shown could be justified by methodological differences including sample collection and preparation, rigor of protein identification and annotation methods. The

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

methodology employed here is tailored to specifically study ECM proteins, while other studies employed more generic proteomics pipeline. In addition, ECM proteins, and collagens in particular, present a specific set of post-translational modifications, namely hydroxylation of lysines and prolines that must be allowed as variable modifications, if, one wants to identify collagen peptides in mass spectrometry datasets (Naba et al., 2017; Taha & Naba, 2019). Of note, our starting material was the cervical region of the root dentin, whereas previous studies utilized the entire root (Jágr et al., 2012; Park et al., 2009). Structural and chemical differences have been identified in different regions of the root (Balooch et al., 2001; Koester et al., 2008), thus regional compositional variations is expected.

The comparison of the young and old human dentin root matrisome identified 2 proteins exclusively in young samples and 4 in old samples, all of which are of interest due to their potential roles in aging. The insulin-like growth factor binding protein 3 (IGFBP3) and lysyl oxidase (LOX) were uniquely identified in young samples. IGFBP3 is an insulin-like growth factor binding protein and recognized to function in the circulation (Baxter, 2000), with age, studies have shown that IGFBP3 appears to be decreased (Blaney Davidson et al., 2005; Hong & Kim, 2018). LOX is an enzyme that has an active role in initiating covalent cross-link formation in collagen fibrils (Kagan & Li, 2003). Its expression is associated to advanced glycation end products (AGEs), age-related modifications to bone mineralization and differentiation of cells in dental tissues (Aoki et al., 2013; Kim et al., 2012). The 4 proteins uniquely identified in old samples are hemopexin (HPX), SERPINC1, S100A9 and transforming growth factor beta-2 (TGFB2). S100A9 is a calcium-binding protein (Zreiqat et al., 2007) implicated in receptors for AGEs (Hiroshima et al., 2018; Xu et al., 2013), alveolar bone resorption and a mediator of aging-associated disorders (Maekawa et al., 2019; Swindell et al., 2013). TGFβ2 is a ligand of the transforming growth factor-beta superfamily (Baxter, 2000) and disruptions of the TGFβ/SMAD pathway has also been linked to disease and age-associated disorders (Tominaga & Suzuki, 2019). HPX and SERPINC1, proteins linked to the circulation and recently associated to Alzheimer's disease (Ashraf et al., 2020) and aging of the skin (Ma et al., 2020), but whose roles in the dentin remain to be understood.

Previous metabolic and proteomic investigations in different model organisms such as mice (Angelidis et al., 2019), nematodes (Copes et al., 2015), and humans (Chaleckis et al., 2016), have shown alterations in ECM abundance with age, however, the underlying mechanisms are yet to be determined. In addition, shifts in protein solubility with age have been shown in protein profiles

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

of young and old superficial digital flexor tendons (Peffers et al., 2014) and between lung samples of young and old mice (Angelidis et al., 2019). However, the functional implications and cause of these changes are still unknown. We suggest that changes in the composition of dentin ECM with age, regarding proteins that are linked to mineralization, collagen post-translational modifications and the accumulation of AGEs could potentially play important roles in alteration of the dentin ECM ultrastructure, mineral and organic content, biomechanical performance and consequently the functionality of teeth with age. Further proteomic investigations using label-based quantitative methods, or focusing on post-translational modifications, and the functional roles of specific proteins may reveal additional differences occurring in the dentin ECM during aging.

5. Conclusions

This study provides a robust experimental framework to explore the composition of the root dentin ECM. Proteomic analysis from single dental roots revealed a feasible profiling of the dentin ECM with limited inter-individual variability among young samples. The analysis of the matrisomes of young and old dentin root samples showed age-related changes to the solubility and composition of the dentin ECM. Future investigations are now necessary to unveil age-dependent mechanisms leading to these changes. In addition, beyond protein identification, we demonstrate the potential of proteomics into gaining structural insight on dentin ECM proteins. With the development of refined methods capable of revealing quantitative (*e.g.*, TMT-labeling) and structural (*e.g.*, degradomics, peptide and PTM mapping on 3D protein structures) information, this proteomic pipeline can become a method of choice for future studies focused on the matrisome of the tooth-periodontal-bone complex.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

CRediT author statement

Mariana Reis: Methodology, Validation, Formal analysis, Investigation, Writing - Review & Editing.

Fred Lee: Methodology, Validation, Formal analysis, Investigation, Writing - Review & Editing.

Ana K. Bedran-Russo: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition.

Alexandra Naba: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition.

Acknowledgements

The authors would like to thank Dr. Hui Chen from the Mass Spectrometry Core facility at the University of Illinois at Chicago and the Dr. George Chlipala from the Research Informatics Core facility at the University of Illinois at Chicago for their technical assistance.

Sources of funding

This project was supported in part by a start-up fund from the Department of Physiology and Biophysics and a Catalyst Award from the Chicago Biomedical Consortium with support from the Searle Funds at the Chicago Community Trust (C-088) to AN and by UIC Wach Research Award to AKBR. Proteomics services were provided by the UIC Research Resources Center Mass spectrometry Core which was established in part by a grant from The Searle Funds at the Chicago Community Trust to the Chicago Biomedical Consortium. Bioinformatic analyses were performed by the UIC Research Informatics Core, supported in part by the National Center for Advancing Translational Sciences (NCATS, Grant UL1TR002003).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Supplementary data

Supplementary Table 1. Complete mass spectrometry dataset

Supplementary Table 2. Comparison of the matrisome of young dentin samples from multiple studies

Supplementary Figure 1. Comparison of the matrisome of young dentin samples from multiple studies

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

Abbey, S. R., Eckhard, U., Solis, N., Marino, G., Matthew, I., & Overall, C. M. (2018). The Human Odontoblast Cell Layer and Dental Pulp Proteomes and N-Terminomes. *Journal of Dental Research*, *97*(3), 338–346. https://doi.org/10.1177/0022034517736054

Aizawa, C., Saito, K., & Ohshima, H. (2019). Regulation of IGF-I by IGFBP3 and IGFBP5 during odontoblast differentiation in mice. *Journal of Oral Biosciences*, *61*(3), 157–162. https://doi.org/10.1016/j.job.2019.07.001

Alford, A. I., Kozloff, K. M., & Hankenson, K. D. (2015). Extracellular matrix networks in bone remodeling. *The International Journal of Biochemistry & Cell Biology*, *65*, 20–31. https://doi.org/10.1016/j.biocel.2015.05.008

Almeida, M., Laurent, M. R., Dubois, V., Claessens, F., O'Brien, C. A., Bouillon, R., Vanderschueren, D., & Manolagas, S. C. (2017). Estrogens and Androgens in Skeletal Physiology and Pathophysiology. *Physiological Reviews*, *97*(1), 135–187. https://doi.org/10.1152/physrev.00033.2015

Angelidis, I., Simon, L. M., Fernandez, I. E., Strunz, M., Mayr, C. H., Greiffo, F. R., Tsitsiridis, G., Ansari, M., Graf, E., Strom, T.-M., Nagendran, M., Desai, T., Eickelberg, O., Mann, M., Theis, F. J., & Schiller, H. B. (2019). An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nature Communications*, *10*(1), 963. https://doi.org/10.1038/s41467-019-08831-9

Aoki, C., Uto, K., Honda, K., Kato, Y., & Oda, H. (2013). Advanced glycation end products suppress lysyl oxidase and induce bone collagen degradation in a rat model of renal osteodystrophy. *Laboratory Investigation*, *93*(11), 1170–1183. https://doi.org/10.1038/labinvest.2013.105

Arola, D., & Reprogel, R. K. (2005). Effects of aging on the mechanical behavior of human dentin. *Biomaterials*, *26*(18), 4051–4061. https://doi.org/10.1016/j.biomaterials.2004.10.029

Ashraf, A., Ashton, N. J., Chatterjee, P., Goozee, K., Shen, K., Fripp, J., Ames, D., Rowe, C., Masters, C. L., Villemagne, V., Hye, A., Martins, R. N., So, P.-W., & AIBL. (2020). Plasma transferrin and hemopexin are associated with altered Aβ uptake and cognitive decline in Alzheimer's disease pathology. *Alzheimer's Research & Therapy*, *12*(1), 72. https://doi.org/10.1186/s13195-020-00634-1

Balasse, M., Bocherens, H., & Mariotti, A. (1999). Intra-bone Variability of Collagen and Apatite Isotopic Composition Used as Evidence of a Change of Diet. *Journal of Archaeological Science*, *26*(6), 593–598. https://doi.org/10.1006/jasc.1998.0376

Balooch, M., Demos, S. G., Kinney, J. H., Marshall, G. W., Balooch, G., & Marshall, S. J. (2001). Local mechanical and optical properties of normal and transparent root dentin. *Journal of Materials Science: Materials in Medicine*, *12*(6), 507–514. https://doi.org/10.1023/A:1011215628779

Barrett, A. S., Maller, O., Pickup, M. W., Weaver, V. M., & Hansen, K. C. (2018). Compartment resolved proteomics reveals a dynamic matrisome in a biomechanically driven model of

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

pancreatic ductal adenocarcinoma. *Journal of Immunology and Regenerative Medicine*, *1*, 67–75. https://doi.org/10.1016/j.regen.2018.03.002

Baxter, R. C. (2000). Insulin-like growth factor (IGF)-binding proteins: Interactions with IGFs and intrinsic bioactivities. *American Journal of Physiology-Endocrinology and Metabolism*, *278*(6), E967–E976. https://doi.org/10.1152/ajpendo.2000.278.6.E967

Blaney Davidson, E., Scharstuhl, A., Vitters, E., van der Kraan, P., & van den Berg, W. (2005). Reduced transforming growth factor-beta signaling in cartilage of old mice: Role in impaired repair capacity. *Arthritis Research & Therapy*, *7*(6), R1338. https://doi.org/10.1186/ar1833

Calderón-González, K. G., Hernández-Monge, J., Herrera-Aguirre, M. E., & Luna-Arias, J. P. (2016). Bioinformatics Tools for Proteomics Data Interpretation. *Advances in Experimental Medicine and Biology*, *919*, 281–341. https://doi.org/10.1007/978-3-319-41448-5_16

Chaleckis, R., Murakami, I., Takada, J., Kondoh, H., & Yanagida, M. (2016). Individual variability in human blood metabolites identifies age-related differences. *Proceedings of the National Academy of Sciences*, *113*(16), 4252–4259. https://doi.org/10.1073/pnas.1603023113

Cho, W. C. S. (2007). Proteomics technologies and challenges. *Genomics, Proteomics & Bioinformatics*, *5*(2), 77–85. https://doi.org/10.1016/S1672-0229(07)60018-7

Chun, S. Y., Lee, H. J., Choi, Y. A., Kim, K. M., Baek, S. H., Park, H. S., Kim, J.-Y., Ahn, J.-M., Cho, J.-Y., Cho, D.-W., Shin, H.-I., & Park, E. K. (2011). Analysis of the Soluble Human Tooth Proteome and Its Ability to Induce Dentin/Tooth Regeneration. *Tissue Engineering Part A*, *17*(1–2), 181–191. https://doi.org/10.1089/ten.tea.2010.0121

Copes, N., Edwards, C., Chaput, D., Saifee, M., Barjuca, I., Nelson, D., Paraggio, A., Saad, P., Lipps, D., Stevens, S. M., & Bradshaw, P. C. (2015). Metabolome and proteome changes with aging in Caenorhabditis elegans. *Experimental Gerontology*, *72*, 67–84. https://doi.org/10.1016/j.exger.2015.09.013

Deutsch, E. W., Bandeira, N., Sharma, V., Perez-Riverol, Y., Carver, J. J., Kundu, D. J., García-Seisdedos, D., Jarnuczak, A. F., Hewapathirana, S., Pullman, B. S., Wertz, J., Sun, Z., Kawano, S., Okuda, S., Watanabe, Y., Hermjakob, H., MacLean, B., MacCoss, M. J., Zhu, Y., … Vizcaíno, J. A. (2020). The ProteomeXchange consortium in 2020: Enabling 'big data' approaches in proteomics. *Nucleic Acids Research*, *48*(D1), D1145–D1152. https://doi.org/10.1093/nar/gkz984

Eckersley, A., Mellody, K. T., Pilkington, S., Griffiths, C. E. M., Watson, R. E. B., O'Cualain, R., Baldock, C., Knight, D., & Sherratt, M. J. (2018). Structural and compositional diversity of fibrillin microfibrils in human tissues. *Journal of Biological Chemistry*, *293*(14), 5117–5133. https://doi.org/10.1074/jbc.RA117.001483

Eckersley, A., Ozols, M., O'Cualain, R., Keevill, E.-J., Foster, A., Pilkington, S., Knight, D., Griffiths, C. E. M., Watson, R. E. B., & Sherratt, M. J. (2020). Proteomic fingerprints of damage in extracellular matrix assemblies. *Matrix Biology Plus*, *5*, 100027. https://doi.org/10.1016/j.mbplus.2020.100027

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Eckhard, U., Marino, G., Abbey, S. R., Tharmarajah, G., Matthew, I., & Overall, C. M. (2015). The Human Dental Pulp Proteome and N-Terminome: Levering the Unexplored Potential of Semitryptic Peptides Enriched by TAILS to Identify Missing Proteins in the Human Proteome Project in Underexplored Tissues. *Journal of Proteome Research*, *14*(9), 3568–3582. https://doi.org/10.1021/acs.jproteome.5b00579

Friedman, P. K., Kaufman, L. B., & Karpas, S. L. (2014). Oral Health Disparity in Older Adults. *Dental Clinics of North America*, *58*(4), 757–770. https://doi.org/10.1016/j.cden.2014.06.004

Gebauer, J. M., & Naba, A. (2020). The Matrisome of Model Organisms: From In-Silico Prediction to Big-Data Annotation. In S. Ricard-Blum (Ed.), *Extracellular Matrix Omics* (Vol. 7, pp. 17–42). Springer International Publishing. https://doi.org/10.1007/978-3-030-58330-9_2

Green, D. R., Schulte, F., Lee, K.-H., Pugach, M. K., Hardt, M., & Bidlack, F. B. (2019). Mapping the Tooth Enamel Proteome and Amelogenin Phosphorylation Onto Mineralizing Porcine Tooth Crowns. *Frontiers in Physiology*, *10*, 925. https://doi.org/10.3389/fphys.2019.00925

Grilo, G. A., Shaver, P. R., Stoffel, H. J., Morrow, C. A., Johnson, O. T., Iyer, R. P., & de Castro Brás, L. E. (2020). Age- and sex-dependent differences in extracellular matrix metabolism associate with cardiac functional and structural changes. *Journal of Molecular and Cellular Cardiology*, *139*, 62–74. https://doi.org/10.1016/j.yjmcc.2020.01.005

Haruyama, N., Sreenath, T. L., Suzuki, S., Yao, X., Wang, Z., Wang, Y., Honeycutt, C., Iozzo, R. V., Young, M. F., & Kulkarni, A. B. (2009). Genetic evidence for key roles of decorin and biglycan in dentin mineralization. *Matrix Biology*, *28*(3), 129–136. https://doi.org/10.1016/j.matbio.2009.01.005

Hayes, M., Burke, F., & Allen, P. F. (2017). Incidence, Prevalence and Global Distribution of Root Caries. In M. Rocha de Olivera Carrilho (Ed.), *Monographs in Oral Science* (Vol. 26, pp. 1–8). S. Karger AG. https://doi.org/10.1159/000479301

Hiroshima, Y., Sakamoto, E., Yoshida, K., Abe, K., Naruishi, K., Yamamoto, T., Shinohara, Y., Kido, J., & Geczy, C. L. (2018). Advanced glycation end-products and Porphyromonas gingivalis lipopolysaccharide increase calprotectin expression in human gingival epithelial cells. *Journal of Cellular Biochemistry*, *119*(2), 1591–1603. https://doi.org/10.1002/jcb.26319

Hong, S., & Kim, M.-M. (2018). IGFBP-3 plays an important role in senescence as an aging marker. *Environmental Toxicology and Pharmacology*, *59*, 138–145. https://doi.org/10.1016/j.etap.2018.03.014

Horsophonphong, S., Sercia, A., França, C. M., Tahayeri, A., Reddy, A. P., Wilmarth, P. A., Surarit, R., Smith, A. J., Ferracane, J. L., & Bertassoni, L. E. (2020). Equivalence of human and bovine dentin matrix molecules for dental pulp regeneration: Proteomic analysis and biological function. *Archives of Oral Biology*, *119*, 104888. https://doi.org/10.1016/j.archoralbio.2020.104888

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Hu, J. C.-C., Yamakoshi, Y., Yamakoshi, F., Krebsbach, P. H., & Simmer, J. P. (2005). Proteomics and genetics of dental enamel. *Cells, Tissues, Organs*, *181*(3–4), 219–231. https://doi.org/10.1159/000091383

Hynes, R. O., & Naba, A. (2012). Overview of the Matrisome -An Inventory of Extracellular Matrix Constituents and Functions. *Cold Spring Harbor Perspectives in Biology*, *4*(1), a004903–a004903. https://doi.org/10.1101/cshperspect.a004903

Iezzi, I., Pagella, P., Mattioli-Belmonte, M., & Mitsiadis, T. (2019). The effects of ageing on dental pulp stem cells, the tooth longevity elixir. *European Cells and Materials*, *37*, 175–185. https://doi.org/10.22203/eCM.v037a11

Iozzo, R. V., & Schaefer, L. (2015). Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix Biology*, *42*, 11–55. https://doi.org/10.1016/j.matbio.2015.02.003

Jágr, M., Eckhardt, A., Pataridis, S., Foltán, R., Myšák, J., & Mikšík, I. (2016). Proteomic analysis of human tooth pulp proteomes – Comparison of caries-resistant and caries-susceptible persons. *Journal of Proteomics*, *145*, 127–136. https://doi.org/10.1016/j.jprot.2016.04.022

Jágr, M., Eckhardt, A., Pataridis, S., & Mikšík, I. (2012). Comprehensive proteomic analysis of human dentin. *European Journal of Oral Sciences*, *120*(4), 259–268. https://doi.org/10.1111/j.1600-0722.2012.00977.x

Jágr, M., Ergang, P., Pataridis, S., Kolrosová, M., Bartoš, M., & Mikšík, I. (2019). Proteomic analysis of dentin-enamel junction and adjacent protein-containing enamel matrix layer of healthy human molar teeth. *European Journal of Oral Sciences*, *127*(2), 112–121. https://doi.org/10.1111/eos.12594

Kagan, H. M., & Li, W. (2003). Lysyl oxidase: Properties, specificity, and biological roles inside and outside of the cell. *Journal of Cellular Biochemistry*, *88*(4), 660–672. https://doi.org/10.1002/jcb.10413

Kawasaki, K., Suzuki, T., & Weiss, K. M. (2004). Genetic basis for the evolution of vertebrate mineralized tissue. *Proceedings of the National Academy of Sciences*, *101*(31), 11356–11361. https://doi.org/10.1073/pnas.0404279101

Keller, A., Nesvizhskii, A. I., Kolker, E., & Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical Chemistry*, *74*(20), 5383–5392. https://doi.org/10.1021/ac025747h

Kim, E.-C., Lee, H.-J., & Kim, Y. (2012). Lysyl Oxidase and the Lysyl Oxidase–like Protein Modulate Odontoblastic Differentiation of Human Dental Pulp Cells. *Journal of Endodontics*, *38*(6), 769–773. https://doi.org/10.1016/j.joen.2012.03.014

Kinney, J. H., Nalla, R. K., Pople, J. A., Breunig, T. M., & Ritchie, R. O. (2005). Age-related transparent root dentin: Mineral concentration, crystallite size, and mechanical properties. *Biomaterials*, *26*(16), 3363–3376. https://doi.org/10.1016/j.biomaterials.2004.09.004

Koester, K. J., Ager, J. W., & Ritchie, R. O. (2008). The effect of aging on crack-growth resistance and toughening mechanisms in human dentin. *Biomaterials*, *29*(10), 1318–1328. https://doi.org/10.1016/j.biomaterials.2007.12.008

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Lau, W. W., Hardt, M., Zhang, Y. H., Freire, M., & Ruhl, S. (2021). The Human Salivary Proteome Wiki: A Community-Driven Research Platform. *Journal of Dental Research*, 002203452110144. https://doi.org/10.1177/00220345211014432

Linde, A. (1989). Dentin matrix proteins: Composition and possible functions in calcification. *The Anatomical Record*, *224*(2), 154–166. https://doi.org/10.1002/ar.1092240206

Loffek, S., Schilling, O., & Franzke, C.-W. (2011). Biological role of matrix metalloproteinases: A critical balance. *European Respiratory Journal*, *38*(1), 191–208. https://doi.org/10.1183/09031936.00146510

Ma, J., Liu, M., Wang, Y., Xin, C., Zhang, H., Chen, S., Zheng, X., Zhang, X., Xiao, F., & Yang, S. (2020). Quantitative proteomics analysis of young and elderly skin with DIA mass spectrometry reveals new skin aging-related proteins. *Aging*, *12*(13), 13529–13554. https://doi.org/10.18632/aging.103461

Madzharova, E., Sabino, F., & Auf dem Keller, U. (2019). Exploring Extracellular Matrix Degradomes by TMT-TAILS N-Terminomics. *Methods in Molecular Biology (Clifton, N.J.)*, *1944*, 115–126. https://doi.org/10.1007/978-1-4939-9095-5_8

Maekawa, S., Onizuka, S., Katagiri, S., Hatasa, M., Ohsugi, Y., Sasaki, N., Watanabe, K., Ohtsu, A., Komazaki, R., Ogura, K., Miyoshi-Akiyama, T., Iwata, T., Nitta, H., & Izumi, Y. (2019). RNA sequencing for ligature induced periodontitis in mice revealed important role of S100A8 and S100A9 for periodontal destruction. *Scientific Reports*, *9*(1), 14663. https://doi.org/10.1038/s41598-019-50959-7

Marshall, G. W. (1993). Dentin: Microstructure and characterization. *Quintessence International (Berlin, Germany: 1985)*, *24*(9), 606–617.

Martin-De Las Heras, S., Valenzuela, A., & Overall, C. M. (2000). The matrix metalloproteinase gelatinase A in human dentine. *Archives of Oral Biology*, *45*(9), 757–765. https://doi.org/10.1016/S0003-9969(00)00052-2

Mukherjee, A., & Rotwein, P. (2008). Insulin-Like Growth Factor-Binding Protein-5 Inhibits Osteoblast Differentiation and Skeletal Growth by Blocking Insulin-Like Growth Factor Actions. *Molecular Endocrinology*, *22*(5), 1238–1250. https://doi.org/10.1210/me.2008-0001

Naba, A., Clauser, K. R., Ding, H., Whittaker, C. A., Carr, S. A., & Hynes, R. O. (2016). The extracellular matrix: Tools and insights for the "omics" era. *Matrix Biology*, *49*, 10–24. https://doi.org/10.1016/j.matbio.2015.06.003

Naba, A., Clauser, K. R., Hoersch, S., Liu, H., Carr, S. A., & Hynes, R. O. (2012). The matrisome: In silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Molecular & Cellular Proteomics*, *11*(4), M111.014647. https://doi.org/10.1074/mcp.M111.014647

Naba, A., Pearce, O. M. T., Del Rosario, A., Ma, D., Ding, H., Rajeeve, V., Cutillas, P. R., Balkwill, F. R., & Hynes, R. O. (2017). Characterization of the extracellular matrix of normal and diseased tissues using proteomics. *Journal of Proteome Research*, *16*(8), 3083–3091. https://doi.org/10.1021/acs.jproteome.7b00191

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Naba, A., & Ricard-Blum, S. (2020). The Extracellular Matrix Goes -Omics: Resources and Tools. In S. Ricard-Blum (Ed.), *Extracellular Matrix Omics* (Vol. 7, pp. 1–16). Springer International Publishing. https://doi.org/10.1007/978-3-030-58330-9_1

Nesvizhskii, A. I., Keller, A., Kolker, E., & Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Analytical Chemistry*, *75*(17), 4646–4658. https://doi.org/10.1021/ac0341261

Pääkkönen, V., Ohlmeier, S., Bergmann, U., Larmas, M., Salo, T., & Tjäderhane, L. (2005). Analysis of gene and protein expression in healthy and carious tooth pulp with cDNA microarray and two-dimensional gel electrophoresis. *European Journal of Oral Sciences*, *113*(5), 369–379. https://doi.org/10.1111/j.1600-0722.2005.00237.x

Park, E.-S., Cho, H.-S., Kwon, T.-G., Jang, S.-N., Lee, S.-H., An, C.-H., Shin, H.-I., Kim, J.-Y., & Cho, J.-Y. (2009). Proteomics Analysis of Human Dentin Reveals Distinct Protein Expression Profiles. *Journal of Proteome Research*, *8*(3), 1338–1346. https://doi.org/10.1021/pr801065s

Pashley, D. H. (1996). Dynamics of the Pulpo-Dentin Complex. *Critical Reviews in Oral Biology & Medicine*, *7*(2), 104–133. https://doi.org/10.1177/10454411960070020101

Peffers, M. J., Thorpe, C. T., Collins, J. A., Eong, R., Wei, T. K. J., Screen, H. R. C., & Clegg, P. D. (2014). Proteomic Analysis Reveals Age-related Changes in Tendon Matrix Composition, with Age- and Injury-specific Matrix Fragmentation. *Journal of Biological Chemistry*, *289*(37), 25867–25878. https://doi.org/10.1074/jbc.M114.566554

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., Pérez, E., Uszkoreit, J., Pfeuffer, J., Sachsenberg, T., Yılmaz, Ş., Tiwary, S., Cox, J., Audain, E., Walzer, M., … Vizcaíno, J. A. (2019). The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Research*, *47*(D1), D442–D450. https://doi.org/10.1093/nar/gky1106

Randles, M., & Lennon, R. (2015). Applying Proteomics to Investigate Extracellular Matrix in Health and Disease. *Current Topics in Membranes*, *76*, 171–196. https://doi.org/10.1016/bs.ctm.2015.06.001

Ravindran, S., & George, A. (2014). Multifunctional ECM proteins in bone and teeth. *Experimental Cell Research*, *325*(2), 148–154. https://doi.org/10.1016/j.yexcr.2014.01.018

Ritchie, H. (2018). The functional significance of dentin sialoprotein-phosphophoryn and dentin sialoprotein. *International Journal of Oral Science*, *10*(4), 31. https://doi.org/10.1038/s41368-018-0035-9

Robinson, C., & Kirkham, J. (1990). The Effect of Fluoride on the Developing Mineralized Tissues. *Journal of Dental Research*, *69*(2_suppl), 685–691. https://doi.org/10.1177/00220345900690S134

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Rogers, L. D., & Overall, C. M. (2013). Proteolytic Post-translational Modification of Proteins: Proteomic Tools and Methodology. *Molecular & Cellular Proteomics*, *12*(12), 3532–3542. https://doi.org/10.1074/mcp.M113.031310

Rosa, N., Correia, M. J., Arrais, J. P., Lopes, P., Melo, J., Oliveira, J. L., & Barros, M. (2012). From the salivary proteome to the OralOme: Comprehensive molecular oral biology. *Archives of Oral Biology*, *57*(7), 853–864. https://doi.org/10.1016/j.archoralbio.2011.12.010

Ryou, H., Romberg, E., Pashley, D. H., Tay, F. R., & Arola, D. (2015). Importance of age on the dynamic mechanical behavior of intertubular and peritubular dentin. *Journal of the Mechanical Behavior of Biomedical Materials*, *42*, 229–242. https://doi.org/10.1016/j.jmbbm.2014.11.021

Salmon, C. R., Tomazela, D. M., Ruiz, K. G. S., Foster, B. L., Paes Leme, A. F., Sallum, E. A., Somerman, M. J., & Nociti, F. H. (2013). Proteomic analysis of human dental cementum and alveolar bone. *Journal of Proteomics*, *91*, 544–555. https://doi.org/10.1016/j.jprot.2013.08.016

Schiller, H. B., Fernandez, I. E., Burgstaller, G., Schaab, C., Scheltema, R. A., Schwarzmayr, T., Strom, T. M., Eickelberg, O., & Mann, M. (2015). Time- and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair. *Molecular Systems Biology*, *11*(7), 819. https://doi.org/10.15252/msb.20156123

Schiller, H. B., Mayr, C. H., Leuschner, G., Strunz, M., Staab-Weijnitz, C., Preisendörfer, S., Eckes, B., Moinzadeh, P., Krieg, T., Schwartz, D. A., Hatz, R. A., Behr, J., Mann, M., & Eickelberg, O. (2017). Deep Proteome Profiling Reveals Common Prevalence of MZB1-Positive Plasma B Cells in Human Lung and Skin Fibrosis. *American Journal of Respiratory and Critical Care Medicine*, *196*(10), 1298–1310. https://doi.org/10.1164/rccm.201611-2263OC

Schiller, H. B., Montoro, D. T., Simon, L. M., Rawlins, E. L., Meyer, K. B., Strunz, M., Vieira Braga, F. A., Timens, W., Koppelman, G. H., Budinger, G. R. S., Burgess, J. K., Waghray, A., van den Berge, M., Theis, F. J., Regev, A., Kaminski, N., Rajagopal, J., Teichmann, S. A., Misharin, A. V., & Nawijn, M. C. (2019). The Human Lung Cell Atlas: A High-Resolution Reference Map of the Human Lung in Health and Disease. *American Journal of Respiratory Cell and Molecular Biology*, *61*(1), 31–41. https://doi.org/10.1165/rcmb.2018-0416TR

Seibel, M. J. (2005). Biochemical markers of bone turnover: Part I: biochemistry and variability. *The Clinical Biochemist. Reviews*, *26*(4), 97–122.

Sharma, V., Rastogi, S., Kumar Bhati, K., Srinivasan, A., Roychoudhury, A., Nikolajeff, F., & Kumar, S. (2020). Mapping the Inorganic and Proteomic Differences among Different Types of Human Teeth: A Preliminary Compositional Insight. *Biomolecules*, *10*(11), 1540. https://doi.org/10.3390/biom10111540

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Socovich, A. M., & Naba, A. (2019). The cancer matrisome: From comprehensive characterization to biomarker discovery. *Seminars in Cell & Developmental Biology*, *89*, 157–166. https://doi.org/10.1016/j.semcdb.2018.06.005

Swindell, W. R., Johnston, A., Xing, X., Little, A., Robichaud, P., Voorhees, J. J., Fisher, G., & Gudjonsson, J. E. (2013). Robust shifts in S100a9 expression with aging: A novel mechanism for chronic inflammation. *Scientific Reports*, *3*(1), 1215. https://doi.org/10.1038/srep01215

Taha, I. N., & Naba, A. (2019). Exploring the extracellular matrix in health and disease using proteomics. *Essays in Biochemistry*, *63*(3), 417–432. https://doi.org/10.1042/EBC20190001

Tominaga, K., & Suzuki, H. I. (2019). TGF-β Signaling in Cellular Senescence and Aging-Related Pathology. *International Journal of Molecular Sciences*, *20*(20), 5002. https://doi.org/10.3390/ijms20205002

United Nations, Department of Economic and Social Affairs. (2019). *World Population Prospects 2019*. https://population.un.org/wpp/

Ventura, T. M. da S., Cassiano, L. de P. S., Souza e Silva, C. M. de, Taira, E. A., Leite, A. de L., Rios, D., & Buzalaf, M. A. R. (2017). The proteomic profile of the acquired enamel pellicle according to its location in the dental arches. *Archives of Oral Biology*, *79*, 20–29. https://doi.org/10.1016/j.archoralbio.2017.03.001

Wei, X., Wu, L., Ling, J., Liu, L., Liu, S., Liu, W., Li, M., & Xiao, Y. (2008). Differentially Expressed Protein Profile of Human Dental Pulp Cells in the Early Process of Odontoblast-like Differentiation In Vitro. *Journal of Endodontics*, *34*(9), 1077–1084. https://doi.org/10.1016/j.joen.2008.06.014

Widbiller, M., Schweikl, H., Bruckmann, A., Rosendahl, A., Hochmuth, E., Lindner, S. R., Buchalla, W., & Galler, K. M. (2019). Shotgun Proteomics of Human Dentin with Different Prefractionation Methods. *Scientific Reports*, *9*(1), 4457. https://doi.org/10.1038/s41598-019-41144-x

Willis, J. R., & Gabaldón, T. (2020). The Human Oral Microbiome in Health and Disease: From Sequences to Ecosystems. *Microorganisms*, *8*(2), 308. https://doi.org/10.3390/microorganisms8020308

Wojtas, M., Lausch, A. J., & Sone, E. D. (2020). Glycosaminoglycans accelerate biomimetic collagen mineralization in a tissue-based in vitro model. *Proceedings of the National Academy of Sciences*, *117*(23), 12636–12642. https://doi.org/10.1073/pnas.1914899117

Xu, X., Chen, H., Zhu, X., Ma, Y., Liu, Q., Xue, Y., Chu, H., Wu, W., Wang, J., & Zou, H. (2013). S100A9 promotes human lung fibroblast cells activation through receptor for advanced glycation end-product-mediated extracellular-regulated kinase 1/2, mitogen-activated protein-kinase and nuclear factor-κB-dependent pathways: S100A9 promotes lung fibroblast activation. *Clinical & Experimental Immunology*, *173*(3), 523–535. https://doi.org/10.1111/cei.12139

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Zhang, J., Sardana, D., Wong, M. C. M., Leung, K. C. M., & Lo, E. C. M. (2020). Factors Associated with Dental Root Caries: A Systematic Review. *JDR Clinical & Translational Research*, *5*(1), 13–29. https://doi.org/10.1177/2380084419849045

Zreiqat, H., Howlett, C. R., Gronthos, S., Hume, D., & Geczy, C. L. (2007). S100A8/S100A9 and their association with cartilage and bone. *Journal of Molecular Histology*, *38*(5), 381–391. https://doi.org/10.1007/s10735-007-9117-2

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure legends

Figure 1. Protein fractionation pipeline for the cervical root dentin.

Schematic of a sound human molar (EN: enamel, D: dentin, C: cementum, DEJ: dentin-enamel junction, CEJ: cement-enamel junction, and P: pulp) and the type of dentin (root) and area (cervical) used for protein extraction (**A**). Experimental pipeline describing sample preparation, protein extraction, digestion, and mass spectrometry analysis (**B**). Color codes used throughout the figures to identify the different protein extracts and age of the individuals (**C**).

Figure 2. Profiling the young dentin matrisome.

Venn diagrams present the number of matrisome proteins detected in supernatants (combined E- and G-extracts) and D-extracts of young individuals (**A**). Venn diagrams present the inter-individual comparisons of young ECM proteins detected in fractions of different solubility: supernatants (**B**) and D-extracts (**C**). Bar chart shows the distribution of proteins detected in young dentin across the different matrisome categories. Proteins included in the analysis were detected with at least two peptides. The young dentin matrisome is defined as the ensemble of proteins detected in at least one of the fractions and individual (**D**). Venn diagram illustrates the number of shared and distinct proteins between the lists defined as the young dentin matrisome in this study and the re-annotated proteomes of human dentin published by Park et al., 2009 and Jágr et al., 2012 (**E *and Supplementary Table 2***).

Figure 3. Profiling the old dentin matrisome

Venn diagrams present the number of matrisome proteins detected in supernatants (combined E- and G-extracts) and D-extracts in old individuals (**A**). Venn diagrams present the inter-individual comparisons of old ECM proteins detected in fractions of different solubility: supernatants (**B**) and D-extracts (**C**). Bar chart shows the distribution of proteins detected in old dentin across the different matrisome categories. Proteins included in the analysis were detected with at least two peptides. The old dentin matrisome is defined as the ensemble of proteins detected in at least one of the fractions and individual (**D**).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 4. Comparison of the young and old dentin matrisomes

Venn diagram illustrates the number of matrisome proteins detected in either or both age groups. Lists of ECM proteins uniquely identified in young (left panel) and old (right panel) individuals (**A**). Bar chart shows the distribution of proteins detected in young (yellow) and old (grey) dentin matrisomes across the different matrisome categories (**B**). Venn diagrams present the comparisons of young and old ECM proteins detected in fractions of different solubility (supernatants and D-extracts) (**C**).

Figure 5. Sequence coverage map of biglycan in young and old dentin samples

Yellow and gray highlights on the biglycan sequence depict the peptides detected by mass spectrometry in young (**A**) and old (**B**) insoluble ECM dentin samples (D-extract). Red sequence highlights depict peptide differentially detected in young and old samples. Average sequence coverage of biglycan in young insoluble ECM dentin samples (D-extract) reaches 38%, while it averaged 14.7% in old insoluble ECM dentin samples ***(see Supplementary Table 1N)***.

Figure 1.

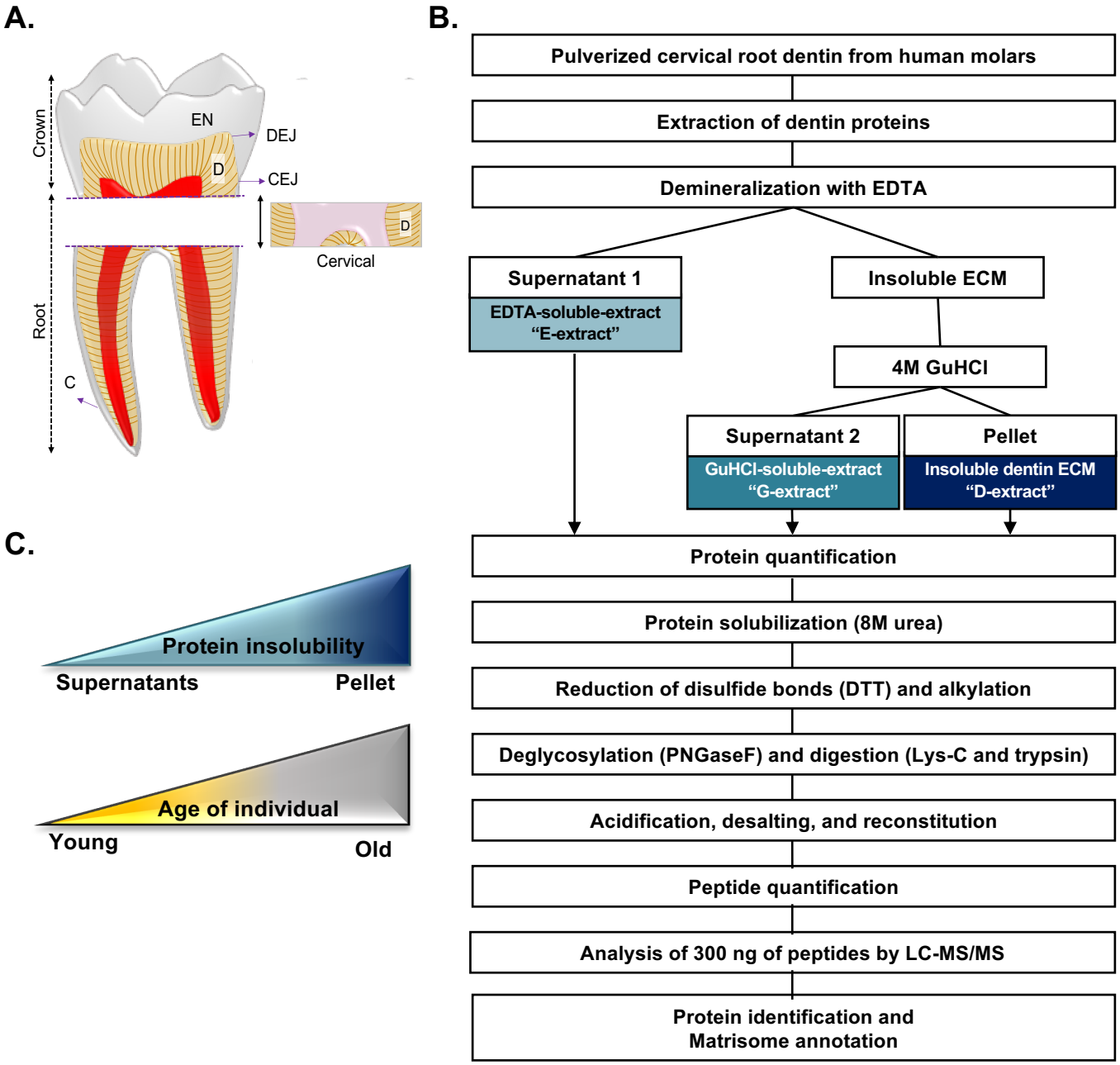

A.
Crown
EN
DEJ
D
CEJ
D
Cervical
Root
C
B.
Pulverized cervical root dentin from human molars
Extraction of dentin proteins
Demineralization with EDTA
Supernatant 1
EDTA-soluble-extract
"E-extract"
Insoluble ECM
4M GuHCl
Supernatant 2
GuHCl-soluble-extract
"G-extract"
Pellet
Insoluble dentin ECM
"D-extract"
Protein quantification
Protein solubilization (8M urea)
Reduction of disulfide bonds (DTT) and alkylation
Deglycosylation (PNGaseF) and digestion (Lys-C and trypsin)
Acidification, desalting, and reconstitution
Peptide quantification
Analysis of 300 ng of peptides by LC-MS/MS
Protein identification and
Matrisome annotation
C.
Protein insolubility
Supernatants
Pellet
Age of individual
Young
Old

Figure 2.

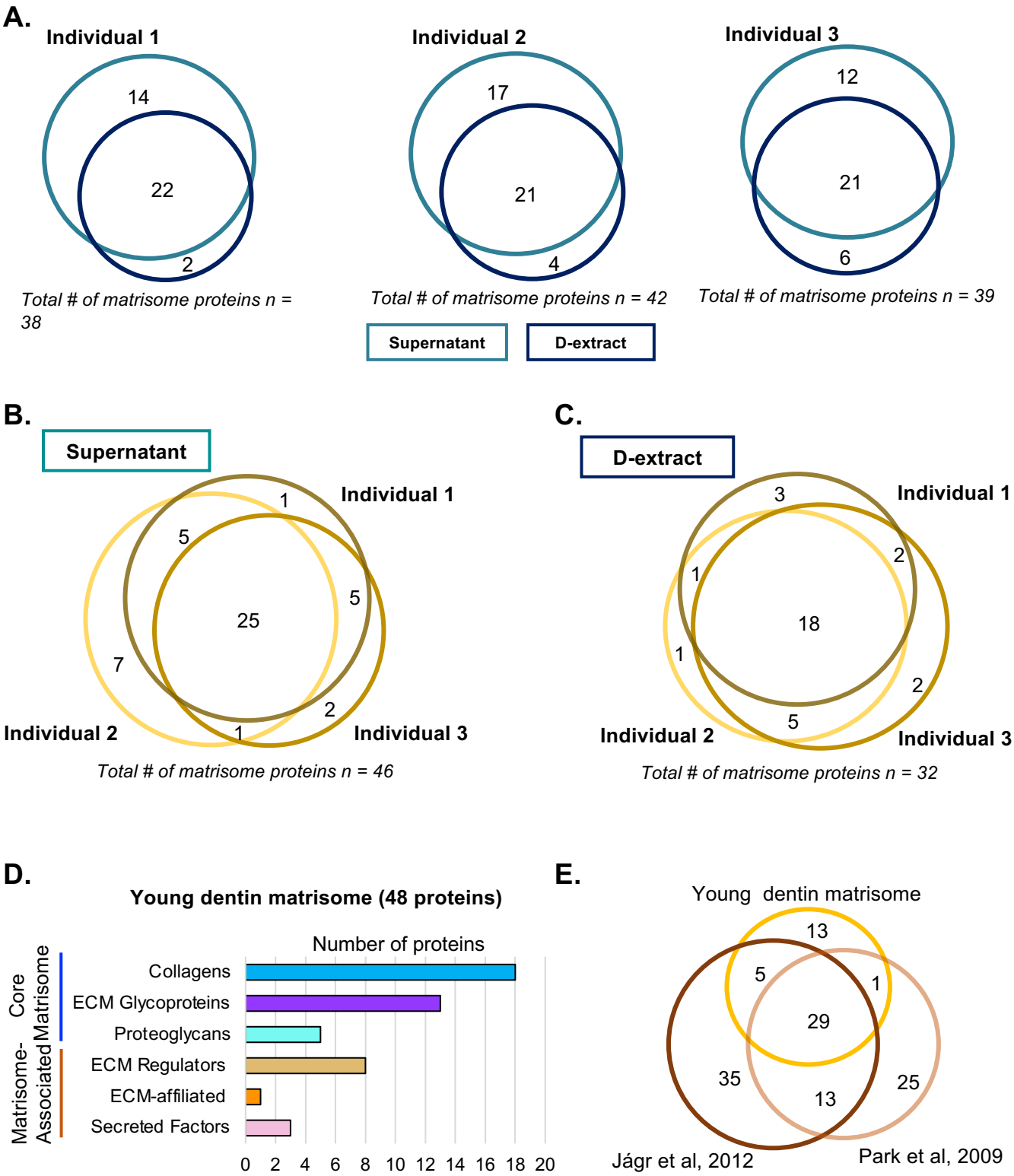
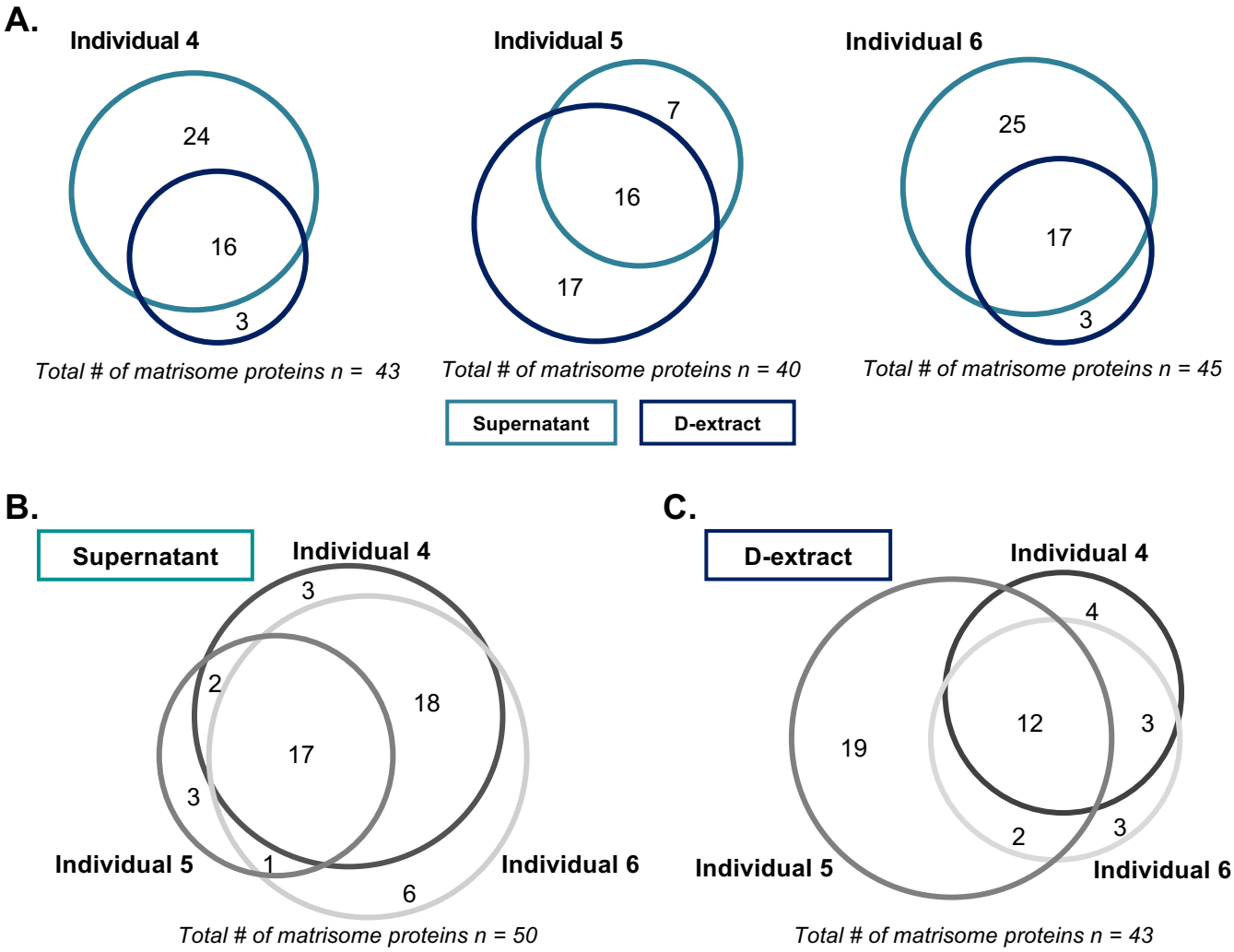

A.
Individual 1
14
22
2
Total # of matrisome proteins n = 38
Individual 2
17
21
4
Total # of matrisome proteins n = 42
Individual 3
12
21
6
Total # of matrisome proteins n = 39
Supernatant
D-extract
B.
Supernatant
Individual 1
1
5
5
25
7
2
1
Individual 2
Individual 3
Total # of matrisome proteins n = 46
C.
D-extract
Individual 1
3
2
1
18
1
2
5
Individual 2
Individual 3
Total # of matrisome proteins n = 32
D.
Young dentin matrisome (48 proteins)
Number of proteins
Core Matrisome
Matrisome-Associated
Collagens
ECM Glycoproteins
Proteoglycans
ECM Regulators
ECM-affiliated
Secreted Factors
0 2 4 6 8 10 12 14 16 18 20
E.
Young dentin matrisome
13
5
1
29
35
13
25
Jágr et al, 2012
Park et al, 2009

Figure 3.


A.
Individual 4
24
16
3
Total # of matrisome proteins n = 43
Individual 5
7
16
17
Total # of matrisome proteins n = 40
Individual 6
25
17
3
Total # of matrisome proteins n = 45
Supernatant
D-extract
B.
Supernatant
Individual 4
3
2
18
17
3
Individual 5
1
Individual 6
6
Total # of matrisome proteins n = 50
C.
D-extract
Individual 4
4
12
3
19
2
3
Individual 5
Individual 6
Total # of matrisome proteins n = 43

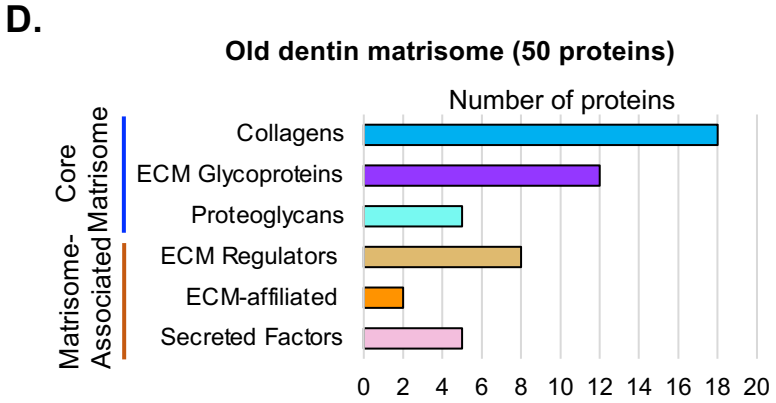
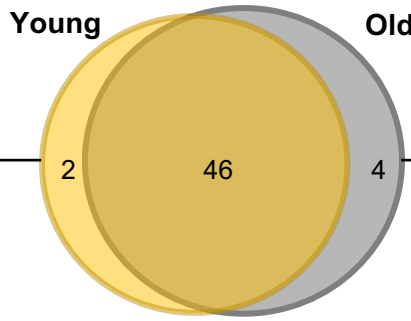

D.
Old dentin matrisome (50 proteins)
Number of proteins
Core Matrisome
Matrisome-Associated
Collagens
ECM Glycoproteins
Proteoglycans
ECM Regulators
ECM-affiliated
Secreted Factors
0 2 4 6 8 10 12 14 16 18 20

Figure 4.

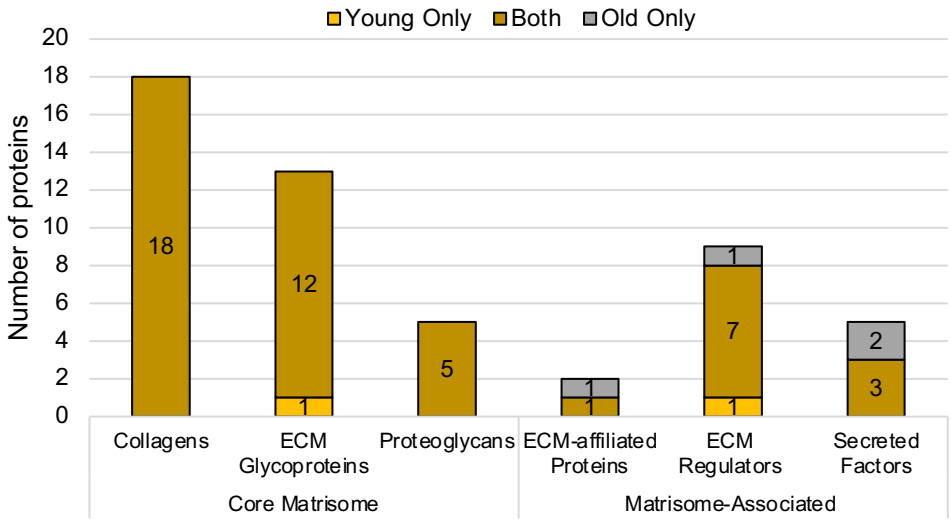
A.

Unique young ECM proteins	
Core Matrisome	Matrisome-Associated
ECM Glycoproteins	ECM Regulators
IGFBP3	LOX

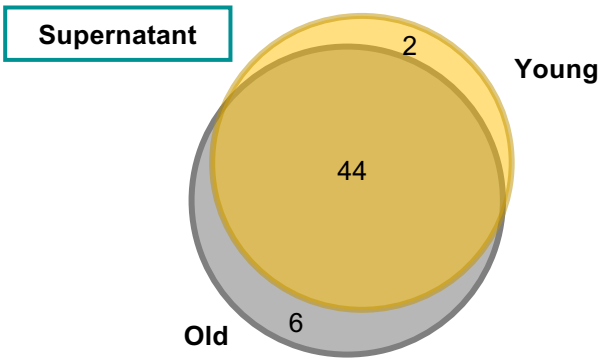


Unique old ECM proteins	
Matrisome-Associated	
ECM-affiliated Proteins	ECM Regulators
HPX	SERPINC1
Secreted Factors	
S100A9	TGFB2

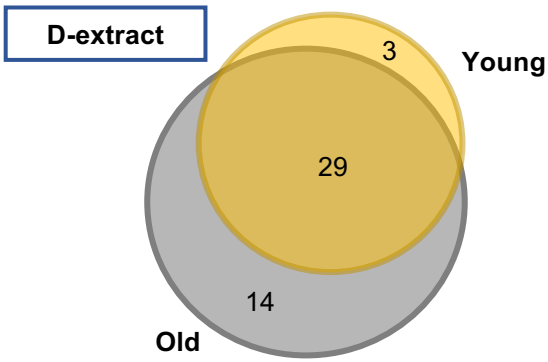
B.



C.



Total # of matrisome proteins n = 52



Total # of matrisome proteins n = 46

Figure 5.

A.

Average sequence coverage of Biglycan in young ECM dentin samples (D-extract): 38%

```
10         20         30         40         50
MWPLWRLVSL LALSQALPFE QRGFWDFTLD DGPFMMNDEE ASGADTSGVL
        60         70         80         90        100
DPDSVTPTYS AMCPFGCHCH LRVVQCSDLG LKSVPKEISP DTTLLDLQNN
       110        120        130        140        150
DISELRKDDF KGLQHLYALV LVNNKISKIH EKAFSPLRKL QKLYISKNHL
       160        170        180        190        200
VEIPPNLPSS LVELRIHDNR IRKVPKGVFS GLRNMNCIEM GGNPLENSGF
       210        220        230        240        250
EPGAFDGLKL NYLRISEAKL TGIPKDLPET LNELHLDHNK IQAIELEDLL
       260        270        280        290        300
RYSKLYRLGL GHNQIRMIEN GSLSFLPTLR ELHLDNNKLA RVPSGLPDLK
       310        320        330        340        350
LLQVVYLHSN NITKVGVNDF CPMGFGVKRA YYNGISLFNN PVPYWEVQPA
       360
TFRCVTDRLA IQFGNYKK
```

B.

Average sequence coverage of Biglycan in old ECM dentin samples (D-extract): 14.7%

```
10         20         30         40         50
MWPLWRLVSL LALSQALPFE QRGFWDFTLD DGPFMMNDEE ASGADTSGVL
        60         70         80         90        100
DPDSVTPTYS AMCPFGCHCH LRVVQCSDLG LKSVPKEISP DTTLLDLQNN
       110        120        130        140        150
DISELRKDDF KGLQHLYALV LVNNKISKIH EKAFSPLRKL QKLYISKNHL
       160        170        180        190        200
VEIPPNLPSS LVELRIHDNR IRKVPKGVFS GLRNMNCIEM GGNPLENSGF
       210        220        230        240        250
EPGAFDGLKL NYLRISEAKL TGIPKDLPET LNELHLDHNK IQAIELEDLL
       260        270        280        290        300
RYSKLYRLGL GHNQIRMIEN GSLSFLPTLR ELHLDNNKLA RVPSGLPDLK
       310        320        330        340        350
LLQVVYLHSN NITKVGVNDF CPMGFGVKRA YYNGISLFNN PVPYWEVQPA
       360
TFRCVTDRLA IQFGNYKK
```

Figure S1.

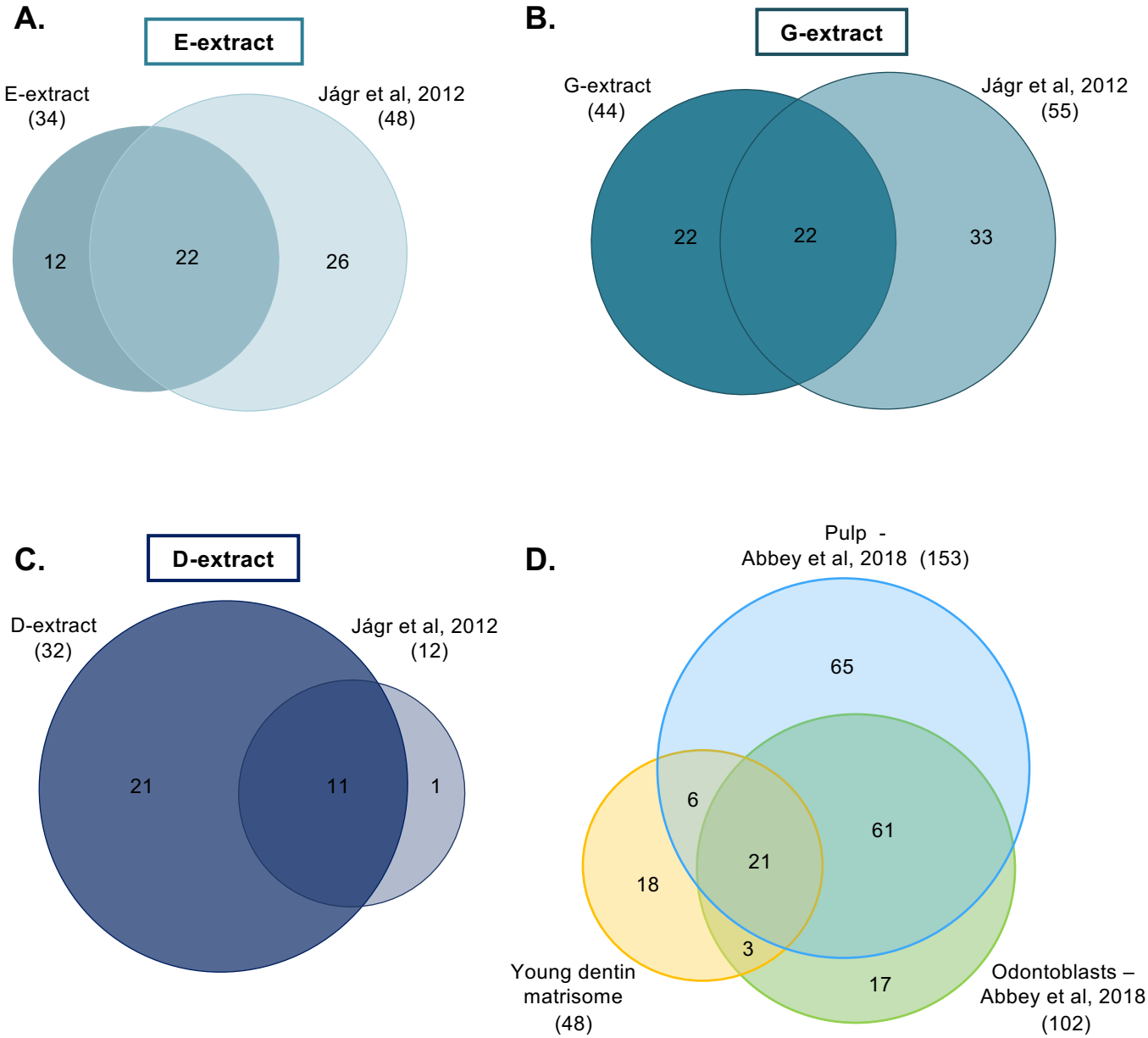


Figure S1. Comparison of the matrisome of young dentin samples from multiple studies. Comparison of the composition of E-extracts **(A)**, G-extracts **(B)**, and D-extracts **(C)**. The complete matrisome of each study is presented in Supporting Table 3, where the proteins are grouped into matrisome divisions and categories. Venn diagrams with the comparison of proteins identified in **(D)** the young dentin matrisome, pulp and odontoblast layer from Abbey et al., 2018. *Related to Supplementary Table 2.*

Table 1. List of ECM proteins of the young dentin matrisome

The proteins included in the list were detected with at least 2 peptides and in one or more individuals. Data are sorted by fractions and matrisome categories (nd: not detected).

Matrisome Division	Matrisome Category	Gene Symbol	Supernatant	D-extract
Core Matrisome	Collagens	COL11A2		
		COL12A1		nd
		COL16A1		
		COL1A1		
		COL1A2		
		COL22A1		
		COL25A1		
		COL28A1		
		COL2A1		
		COL3A1		
		COL4A1		
		COL4A2		nd
		COL4A5		
		COL5A1		
		COL5A2		
		COL5A3		
		COL7A1		
		COL9A3	nd	
	ECM Glyco-proteins	CTHRC1		
		DPT		
		DSPP		nd
		FBN1	nd	
		IGFBP3		nd
		IGFBP5		
		MGP		nd
		PCOLCE		
		POSTN		
		SPARC		nd
		SPP1		nd
		THBS1		nd
		VTN		
	Proteoglycans	ASPN		nd
		BGN		
		CHAD		nd
		LUM		nd
		OMD		
Matrisome Associated	ECM Regulators	F10		
		F2		
		F9		nd
		KNG1		
		LOX		nd
		MMP20		
		SERPINA1		nd
		SERPINF1		
	ECM-affiliated	CLEC3B		
	Secreted Factors	IGF2		nd
		PTN		
		TGFB1		nd

Table 2. List of ECM proteins of the old dentin matrisome

The proteins included in the list were detected with at least 2 peptides and in one or more individuals. Data are sorted by extracts and matrisome categories (nd: not detected).

Matrisome Division	Matrisome Category	Gene Symbol	Supernatant	D-extract
Core Matrisome	Collagens	COL11A2		
		COL12A1		
		COL16A1		
		COL1A1		
		COL1A2		
		COL22A1		
		COL25A1		
		COL28A1		
		COL2A1		
		COL3A1		
		COL4A1		
		COL4A2		
		COL4A5		
		COL5A1		
		COL5A2		
		COL5A3		
		COL7A1		
		COL9A3		
	ECM Glyco-proteins	CTHRC1		
		DPT		nd
		DSPP		
		FBN1		nd
		IGFBP5		
		MGP		
		PCOLCE		
		POSTN		
		SPARC		
		SPP1		nd
		THBS1		
		VTN		
	Proteoglycans	ASPN		
		BGN		
		CHAD		nd
		LUM		
		OMD		
Matrisome Associated	ECM Regulators	F10		
		F2		
		F9		
		KNG1		
		MMP20		
		SERPINA1		
		SERPINC1		
		SERPINF1		
	ECM-affiliated	CLEC3B		
		HPX		nd
	Secreted Factors	IGF2		
		PTN		nd
		S100A9		
		TGFB1		
		TGFB2		nd

Table 3. Comparing the young and old dentin matrisome

The proteins included in the list were detected with at least 2 peptides and in one or more individuals. Data are sorted by extracts and matrisome categories (nd: not detected).

Matrisome Division	Matrisome Category	Gene Symbol	Supernatant	D-extract	Supernatant	D-extract
Core Matrisome	Collagens	COL11A2				
		COL12A1		nd		
		COL16A1				
		COL1A1				
		COL1A2				
		COL22A1				
		COL25A1				
		COL28A1				
		COL2A1				
		COL3A1				
		COL4A1				
		COL4A2		nd		
		COL4A5				
		COL5A1				
		COL5A2				
		COL5A3				
		COL7A1				
		COL9A3	nd			
	ECM Glyco-proteins	CTHRC1				
		DPT				nd
		DSPP		nd		
		FBN1	nd			nd
		IGFBP3		nd	nd	nd
		IGFBP5				
		MGP		nd		
		PCOLCE				
		POSTN				
		SPARC		nd		
		SPP1		nd		nd
		THBS1		nd		
		VTN				
	Proteoglycans	ASPN		nd		
		BGN				
		CHAD		nd		nd
		LUM		nd		
		OMD				
Matrisome Associated	ECM Regulators	F10				
		F2				
		F9		nd		
		KNG1				
		LOX		nd	nd	nd
		MMP20				
		SERPINA1		nd		
		SERPINC1	nd	nd		
		SERPINF1				
	ECM-affiliated	CLEC3B				
		HPX	nd	nd		nd
	Secreted Factors	IGF2		nd		
		PTN				nd
		S100A9	nd	nd		
		TGFB1		nd		
		TGFB2	nd	nd		nd