bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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 Profiling of the Human Dentin Identifies Age-Related Differences in the Composition and Solubility of the Matrisome

Mariana Reis[1,2], Fred Lee[3], Ana K. Bedran-Russo[1,2*], Alexandra Naba[3*]

Author Affiliation:

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

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

[3] 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*

***Keywords*:**

Extracellular matrix, Tooth, Aging, Dentin, Collagens, Proteoglycans, Protein solubility, Mass spectrometry

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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

Aging is the conjunction of progressive physiological changes that occur in living organisms over time. A fundamental component of organs and tissues that undergoes functional and structural modifications with age is the extracellular matrix (ECM). The ECM is a dynamic assembly of proteins that regulates cellular functions and participates in tissue organization and remodeling. The ECM of teeth is no different but lacks turnover capability. Dentin, the bulk of the tooth, endures modifications throughout a lifespan. As of now, the composition of the dentin ECM and how it changes with age is unknown. To address, we used proteomics to profile the dentin ECM of young and older adult human teeth. We report the identification of 341 unique ECM proteins and define the matrisome of young and old dentin as composed of 125 and 112 proteins, respectively. Our study identified over 100 ECM proteins that have not previously been associated with the dentin. Importantly, it also reveals changes to both protein solubility and matrisome composition as a function of age. Our study is a first step towards the discovery of biomarkers of the aging tooth and the development of strategies to maintain or regenerate a healthy dentition in our aging population.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

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[1]. These cohorts will then sustain their natural dentition[2] and experience higher prevalence of oral health conditions, as caries and periodontal disease[3,4]. Despite considerable research in recent decades, the aging process of human teeth remain largely unknown.

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[5–7]. The ECM of bones has modeling and remolding abilities, and experience continuous turnover, via resorption and formation of new tissue orchestrated by osteoclasts and osteocytes[8]; however, unlike bone, teeth lack such turnover[9], as the mineralized tissues are acellular. Odontoblast cells are confined into the pulp chamber and deposit continuously dentin layers over time[10]. The dentin is the major component of teeth and forms a three-dimensional gradual organization of tubules that extends from the pulp chamber to form a dentin-pulp complex (Fig. 1A). The dentin is composed of a collagenous and a non-collagenous ECM reinforced by a carbonated nanocrystalline apatite mineral phase[9–12].

Collagen type I is the primary structural protein found in dentin. Alongside collagens type II-VI, XI and XII, they provide scaffolding, ECM organization, and mechanical properties[13–15]. Phosphoproteins from the SIBLING (Small Integrin-Binding Ligand, N-linked Glycoprotein) family, such as dentin phosphoprotein (DPP), dentin sialoprotein (DSP), dentin sialophosphoprotein (DSPP), dentin matrix protein-1 (DMP1) and osteopontin (OPN), have important regulatory roles in tissue mineralization[16–20]. Glycoproteins, including matrix Gla-proteins (MGP), SPARC, and growth factors (transforming growth factors beta family) regulate signal transduction, cell proliferation, regulation and potentially contribute to dentin repair[16]. Decorin (DCN), biglycan (BGN), lumican (LUM), fibromodulin (FMOD) and osteomodulin (OMD) small leucine-rich proteoglycans (SLRPs)[14,15,21,22] serve regulatory and structural roles including tissue hydration, mineralization processes, space-filling and mechanical support[23–25]. Once the tooth is in function, the dentin ECM components maintain the structural integrity and mechanical features of the tooth, while also slowing the progression of dental caries[10,26]. Of note, until now, investigations on the content and assembly of the dentin ECM have in large utilized young human teeth.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

Dentin undergoes physiological and pathological modifications[10]. With aging, dentin exhibits higher mineral content[27,28], a gradual reduction of dentin tubule diameter and dental pulp size[29], and decreased mechanical properties[30–32]. However, there is a significant gap of knowledge of the composition, abundance, and state of function of the dentin ECM over the lifespan of an individual. Thus, having a cartography of the dentin ECM during aging, would be a first step towards the discovery of biomarkers of the aging tooth and the development of strategies preventing or slowing down the pathological processes associated with the aging human dentition.

Proteomic approaches have previously been used to study human teeth and have used chaotropic agents such as guanidine hydrochloride[33], or detergent-containing solutions like RIPA[14] to identify ECM components of the tooth. Sequential extraction using guanidine hydrochloride and EDTA further increased the number of proteins identified in the dentin by extracting proteins of different solubilities[15]. But, while these studies identified a various number of proteins present in human dentin, they were not designed to investigate specifically ECM proteins. In addition, the absence of using a unified annotation of ECM components prevented a comprehensive definition of the dentin ECM proteins and comparisons across different cohorts.

The large size and high insolubility of ECM proteins have made their biochemical analysis notoriously difficult. However, over the past decade, we and others have developed experimental proteomic pipelines to enrich, solubilize, and digest ECM proteins[34–36]. 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[37]. 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[38], or during aging[39,40]. Here, we employed ECM proteomics to identify age-related changes in the composition and solubility of ECM proteins from the root dentin of young and older adult individuals to aid in building the tooth matrisome.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

Material and methods

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 S1A). Teeth were immediately stored at -20ºC after extraction. 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.

Dentin protein extraction

The extraction of dentin proteins followed an existing protocol[14,33], with modifications (Fig. 1B). In brief, the cervical dentin root sections were pulverized (250-350 mg per tooth/sample) 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. Further protein extraction was performed with a 4M guanidine hydrochloride (Sigma-Aldrich St. Louis, MO, USA)] solution at pH 7.4 containing protease inhibitors, for 3 days under agitation at 4°C. The EDTA and guanidine hydrochloride supernatants, E-extract, and G-extract, respectively, were obtained by centrifugation at 5000 x g for 10 min at 4°C. 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- and G-extracts and dentin pellets were lyophilized (FreeZone Freeze Dryer, Labconco, Kansas City, MO, USA) and stored at -20°C for further analyses.

Protein digestion

The lyophilized samples (E-extract, G-extract and D-extract, 5-10 mg dry weight) were subsequently solubilized and digested into peptides following an established and previously described protocol (Fig. 1B)[41]. 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

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

deglycoslyated with PNGaseF (New England Biolabs, Ipswich, MA, USA) and digested with Lys-C (Thermo Fisher Scientific, Rockford, IL, USA), and trypsin (Thermo Fisher Scientific, Rockford, IL, USA), at 37 °C. The samples were acidified with 50% trifluoroacetic acid (TFA, Thermo Fisher Scientific, Rockford, IL, USA) until the sample pH reached below 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) and approximately 300 ng of peptides were analyzed by LC-MS/MS.

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 (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.

RAW files were converted into mgf files using MSConvert (ProteoWizard). Database search was carried out using Mascot server version 2.6.2 (from Matrix Science). Mascot search parameters were: 10 ppm mass tolerance for precursor ions; 100 mmu for fragment-ion mass tolerance; two missed cleavages of trypsin; fixed modification was carbamidomethylation of cysteine; and variable modifications were oxidized methionine, deamidation of asparagine, pyro-glutamic acid modification at N-terminal glutamine, and hydroxylation of lysine and proline. Only peptides with a Mascot score ≥ 25 and an isolation interference ≤ 30 were included in the data analysis. For our analyses, we included proteins detected with at least 2

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

unique peptides. Mass spectrometry output was further annotated to identify ECM and non-ECM components using the matrisome list we previously[37,41] devised and newly developed R-scripts available on GitHub at https://github.com/uic-ric/matrisome_tools. In brief, matrisome components are classified as core-matrisome or matrisome-associated components, and further categorized into groups based on structural 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[41,42]. Semi-quantitative analysis of proteins found in the D-extract of both young and old samples was conducted by comparing the normalized MS1 intensity. Statistical significance was determined by an unpaired t-test assuming equal variance of the samples (Supplementary Table S1O).

Raw mass spectrometry data have been deposited to the ProteomeXchange Consortium[43] via the PRIDE partner repository[44] with the dataset identifier PXD018320 and 10.6019/PXD018320. *The raw data will be made publicly available upon acceptance of the manuscript.*

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

Results and discussion

Previous studies have shown that dentin proteins are soluble in either EDTA or guanidine solutions[14,15], we thus aimed to analyze the ECM content of two fractions of different solubility (E- and G-extracts) but also of the remaining guanidine-hydrochloride-insoluble dentin pellet (D-extract), so as to obtain the highest amount of protein from the dentin ECM.

In order to assess the reproducibility of the experimental pipeline devised for this study, we first compared, for each individual of the two age groups, the MS1 signal intensity (or peptide abundance) and the numbers of spectra and proteins detected in each of the three extracts of different protein solubility (Fig. 2A, 2B and Supplementary Table S1A). Overall, and as expected, peptides derived from matrisome proteins contributed to the majority of the MS1 signal intensity in all extracts (Fig. 2A, 2B and Supplementary Table S1A). A total of 260 and 247 distinct ECM proteins were detected in young and old individuals, respectively (Supplementary Table S1D and S1I).

2.1. Assessing the inter-individual variability in the ECM protein composition in dentin extracts of different solubility

We first sought to evaluate the inter-individual relationships of E-, G- and D-extracts of for each age group (Fig. 3A and 3B). Our assessment of the ECM protein composition in distinct extracts revealed that inter-individual variability observed was limited and comparable among extracts and age brackets. We also show that our pipeline identifies well over 100 distinct ECM proteins in each condition (*see below*).

By taking into account ECM proteins detected with two or more peptides, our data identify respectively 126, 168, and 161 distinct ECM Proteins in the E-, G- and D-extracts of young individuals (Fig. 3A). Moreover, we identified in young individuals a set of 44 ECM proteins in E-extracts that overlap between all 3 individuals, 53 in G-extracts and 59 in D-extracts. The number of unique ECM proteins found in each young individual varied according to the extract, where a higher number of unique ECM proteins was detected in G-extracts (92), when compared to E- (60) and D-extracts (66) (Fig. 3A), indicating a higher variability in samples of intermediate solubility.

We found that extracts obtained from older individuals showed a total number of ECM proteins of 158 in E-extracts, 127 in G-extracts, and 156 in D-extracts. A greater variability was seen in D-extracts (84), when compared to E- (72) and G-extracts (58) (Fig. 3B). The

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

overlap of ECM proteins between older individuals was similar, where a set of 45, 40, and 47 ECM proteins were detected in E-, G- and D-extracts, respectively (Fig. 3B).

Overall, the inter-individual assessment revealed a similar level of variability between extracts and individuals of different age groups, where 40 to 55% of the ECM proteins identified were individual specific, 14 to 26% were present in two out of three individuals and 28 to 37% were common in all three individuals.

2.2. Dentin ECM protein solubility varies with age

The comparison of the matrisome proteins detected in the extracts representing proteins of different solubility showed that each extract has a different protein composition (Fig. 4A-F).

The number of ECM proteins identified in all three extracts of samples obtained from young individuals represented 40 to 60% of the ECM proteins. The numbers of unique ECM proteins detected in a specific extract were similar among individuals, in which the D-extracts (containing the most insoluble proteins) had the highest number of unique proteins, 35-38%, and the E-extracts (containing more soluble proteins), the lowest, 20-33% (Fig. 4A-C). Interestingly, we observed that, overall, G- and D-extracts of young individuals had similar ECM content (number of proteins and peptide abundance or MS1 intensity), although higher ECM content, when compared to E-extracts (Fig. 2A, 4A-C, and Supplementary Table S1A). This was somewhat anticipated, since ECM proteins, being large and highly cross-linked, are highly insoluble.

In contrast, the analysis of the dentin ECM protein composition of old individuals revealed a higher number of ECM proteins in samples of higher solubility (E-extracts; Fig. 4D-F, and Supplementary Table S1A) and approximately 25% less proteins detected in the G-extracts when compared to the G-extracts of young individuals. This change was also reflected at the peptide abundance (MS1 intensity) level. The number of shared proteins identified in extracts of old individuals represented 37 to 60% of the ECM proteins (Fig. 4D-F). We observed that the number of unique ECM proteins detected in a specific extract had a wider range among individuals, namely 32-44% (E-extract), 19-30% (G-extract) and 27-37% (D-extract).

Altogether, these results suggest a shift in the solubility of ECM protein of dentin with age, where the proteins composing the dentin of older teeth are more soluble than the proteins of the dentin ECM of younger teeth (Fig. 3, lower panels). Multiple mechanisms contribute to

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

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. Previous studies have found shifts in protein solubility with age, one demonstrating differences in protein profiles of young and old superficial digital flexor tendons, and altered ECM protein turnover and increased protein degradation in old samples when compared to young[45]; moreover a more recent investigation showed a comparison of solubility profiles between lung samples of young and old mice, and revealed a variety of proteins with altered solubility profiles as a function of age[40]. However, the functional implications and cause of these changes to protein solubility with aging are still unknown and we anticipate that it will be the focus of future investigations.

2.3. Defining the dentin matrisome

In order to further characterize age-related differences in the ECM composition of young and old dentin, we defined, for each age group, the dentin matrisome as the ensemble of proteins detected in at least two individuals and in either E-, G- and D-extracts (see triangles in Fig. 3A and 3B, and resulting data Fig. 4G and 4H and Table 1). Using this definition, the dentin matrisome of young and old adult individuals comprises, 125 and 112 ECM proteins, respectively (Fig. 4G and 4H, Table 1).

The main proteins that have previously been shown to compose the dentin ECM were found, in majority, in the E- and D-extracts of both young and old samples. Collagens, as expected, were the most abundant protein in the dentin matrisomes. The fibrillar type I collagen (COL1A1 and COL1A2) was predominant[46], however, we also report the identification of a large number of collagens not previously associated with the dentin tissue, including collagens type VII-X, XIII-XIX and XXI-XVIII (Table 1).

We identified glycoproteins such as DSPP, SPP1, PCOLCE, SPARC, MGP and VTN that are known dentin proteins with important roles in biomineralization, enzymatic cleavage of type I procollagen and supporting cell adhesion[16,20,47,48]. In addition, we also report the detection of other glycoproteins never found before in the dentin ECM, such as AMBN, AMELY, COLQ, CRELD1, CTHRC1, ECM1, EMID1, EMILIN2, FBLN2, FNDC1 and FNDC8, FRAS1, GLDN, HMCN2, IGSF10, LAMA2,3 and 5, MXRA5, SSPO and TNXB, (*see below*).

Proteoglycans (PGs) from the small leucine-rich proteoglycan (SLRP) family were the major PGs herein found. They include BGN, DCN, LUM and OMD, which are known to

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

participate in collagen fibril assembly, tissue growth and regulation of interfibrillar spacings[23,49]. We also detected perlecan (HSPG2), one of the largest PGs that binds to different ECM components and plays various roles in cell signaling, mineralization and biomechanics[50].

Among the ECM-associated proteins, we report the identification of large number of A Disintegrin And Metalloproteinase (ADAM) family members, such as ADAMTS12, ADAM2, ADAMTS2, ADAMTSL3, ADAM28, ADAM8, ADAM29, ADAMTSL5 (Table 1). These proteins are involved in proteolytic release of fragments of cell surface proteins[51]. Additionally, we detected ECM-affiliated proteins from the semaphorins protein family, SEMA6D, SEMA4G and SEMA5B, that are known for their effects on cellular processes (adhesion, aggregation, migration, etc.)[52] .

In summary, the depth of our analysis, has permitted the identification of proteins never found before in the dentin ECM and is opening novel avenues for dental research (*see Conclusion section*).

2.4. The dentin ECM composition changes with age

The comparison of the young and old dentin ECM matrisome composition revealed that a large portion of proteins (91) are identified in both age groups (Fig. 5A and Supplementary Table S1N). Remarkably, a higher number of unique ECM proteins was found in young teeth (34) when compared to old teeth (21) (Figure 5A, Table 1). The young dentin matrisome consists of predominantly collagens (34%) and glycoproteins (27%), along with proteoglycans (5%), where together they constitute the young dentin core matrisome. Matrisome-associated proteins were also present, however in a smaller amount, represented by 14% each of ECM-affiliated proteins and regulators, and 7% of secreted factors (Supplementary Table S1H). The old dentin matrisome is composed of a majority of core matrisome proteins, where 38% are collagens, 22% ECM glycoproteins and 4% proteoglycans. The matrisome-associated proteins were less evident shown by 14% of ECM-affiliated proteins, 13% of regulators, and 7% of secreted factors (Supplementary Table S1M).

The major components found in both groups are collagens (46%) and ECM glycoproteins (23%) (Fig. 5B and Table 1). A smaller number of proteoglycans and secreted factors were detected, approximately 4 to 5%, each. Matrisome-associated proteins such as ECM-affiliated proteins and regulators were identified in similar number (10-11%). Overall, the core matrisome proteins were predominant (74%) in the proteins found in both young and old samples.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

Tables in Fig. 5C and 5D show the ECM proteins uniquely identified in the young or old dentin matrisomes. Importantly, we identified significant age-related shifts in the number of unique ECM proteins, shown predominantly by a 3.25-fold increase in the number of ECM glycoproteins in young individuals when compared to older individuals. Exclusively found in the young matrisome were ameloblastin (AMBN) and amelogenin Y-Linked (AMELY), two glycoproteins that primarily have an important role in enamel matrix formation[53]; however, studies have reported the presence of these proteins in other mineralized and soft tissues (i.e. dentin, cementum, periodontal ligament, bone, brain, and soft tissues), where they participate in the mineralization process[54,55]. Other noteworthy changes were to the PGs content, in which decorin (DCN) and neurocan (NCAN) were detected solely in the young matrisome (Fig. 5C), while, PRG4 was uniquely detected in the old dentin matrisome (Fig. 5D). The changes in PGs abundance with age has not been fully explored in dentin, however, a study concerning the aging human skin showed that PGs such as versican and decorin displayed age-related differences, where the size and sulfation pattern of glycosaminoglycans (GAGs) of versican decreased with age, as did the length of the GAGs of decorin[56]. Hence, we suggest that changes in the composition of dentin ECM with age, regarding PGs, could potentially play important roles in alterations of the dentin ECM ultrastructure, collagenous network and eventually biomechanics and functionality.

The number of affiliated proteins and secreted factors were comparable among age groups, yet ECM regulators of young dentin showed 1.6-fold higher number of unique proteins when compared to old samples. The decrease in the total number of ECM regulators in dentin with age, in particular proteinases (ADAMs and ADAMTSs and their inhibitors) could contribute to decreased ECM remodeling, leading to the disorganization of the collagenous ECM[57], and overall structural and regulatory damage to the aging dentin ECM.

The comparison between young and old matrisomes ultimately reveals that, while sharing a large number of proteins, each matrisome presents certain unique components that may relate to the particular functions of the dentin tissue. Previous metabolic and proteomic investigations in different model organisms such as mice[40], nematodes[58], and humans[59], have shown alterations in ECM abundance with age, however, the underlying mechanisms are yet to be determined.

Since our study is the first to report the in-depth characterization of the most insoluble proteins (D-extract) of the dentin in young and old samples, we further analyzed in more details the composition and abundance of D-extracts. Our analysis revealed that a large fraction of matrisome proteins in D-extract (59) are detected in both young and old groups (Fig. 6, Table

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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). The main protein category identified was collagen, represented by 68% of the total proteins. ECM glycoproteins accounted for 14% of the shared proteins, proteoglycans for 3%, and matrisome associated proteins, for 15% total. Interestingly, there were more ECM proteins solely identified in young individuals (36) when compared to old (13) individuals (Fig. 6, Table 1). Predominantly, the unique proteins detected, in young individuals, were ECM glycoproteins (31%), ECM-affiliated proteins (22%) and regulators (25%), whereas in old individuals mainly matrisome-associated proteins (ECM-affiliated, 31% and regulators, 23%) were found.

Semi-quantitative analysis was accomplished by comparing the relative protein abundance in the ECM proteins detected in both young and old D-extracts using MS1 intensity data (Supplementary Table S1O). Only proteins represented with at least two peptides were considered for the analysis. Of all 59 proteins found in both young and old samples (Table 1 and Supplementary Table 1O), only COL3A1 was detected in statistically significant higher abundance in old D-extract when compared to young (Supplementary Table 1O; $p = 0.034$). Collagen type III is a fibrillar collagen that is highly associated with collagen type I[60]. Its deficiency is related with Ehlers-Danlos syndrome (EDS)[61]. COL3A1 has also been associated with regulation of fibril diameter and collagen cross-linking[62], thus we suggest that higher abundance of this protein in older adult dentin ECM could be associated with collagen post-translational modifications that occur with time.

Even though, the abundance of ECM proteins is comparable among young and old dentin, 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.

2.5. Meta-analysis of young dentin proteomic datasets

Last, we aimed to compare the list of ECM proteins defined in our study with those of reported in two previous studies that have explored the proteome of human dentin root[14,15]. In order to compare our data with these, we retrieved and reannotated them with the *in silico*-predicted complete matrisome list[63] (Supplementary Table S2A). Since the two selected studies analyzed dentin samples from individuals between the ages of 10 and 33, the comparison of the published data was performed with the young dentin matrisome defined here (Supplementary Table S2B).

The assessment revealed 27 ECM proteins found in all three studies (COL1A1, COL1A2, COL5A1, COL11A2, COL12A1, DPP, DSPP, BGN, DCN, HSPG2, MMP20, VTN,

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

IGFBP5, MGP, PCOLCE, POSTN, SPARC, SPP1, LUM, OMD, CLEC3B, F2, F10, KNG1, SERPINA1, SERPINA10, SERPINF1) and another set of 24 ECM proteins that were detected in two of the three studies (Fig. 7A, Table 2). The presence of proteins such as DPP, DSPP, BGN, DCN, HSPG2, and 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 22 proteins were unique to the studies by Jágr et al[15]. and Park et al[14]., respectively, our study identified an additional 89 ECM proteins not detected before in the dentin of the root of human teeth. These 89 proteins consisted of 31 collagens, 24 glycoproteins, 16 ECM-affiliated proteins, 9 secreted factors, 8 ECM regulators, and 1 proteoglycan (Table 2, Supplementary Table S2B). This meta-analysis overall demonstrates the depth of our ECM-focused proteomic and analytical pipeline and reveals an extensive list of collagens never before identified in this type of samples. This can be explained by the fact that our methodology was tailored to study specifically ECM proteins, while other studies employed more generic proteomics pipeline, for example, core ECM proteins being typically very large, SDS-PAGE, may not allow to fully resolve them. 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, as we previously reported, one wants to identify collagen peptides in mass spectrometry datasets[36,64].

Since our approach allowed for the specific profiling of three distinct dentin extracts, we further compared extracts of similar solubility from the study from Jágr et al to the respective E-, G-, and D-extracts of the present study, (Fig. 7B, 7C, and 7D, Table 2). A set of 23 ECM proteins were identified in E-extracts of both studies (Fig. 7B). G-extracts and D-extracts revealed 15 and 10 ECM overlapping proteins, respectively (Fig. 7C and 7D). While the comparison of G-extracts revealed that similar number of proteins that were uniquely identified in each study, the D-extract, contained a significantly greater number of ECM proteins identified (85) compared to Jágr et al (2). Altogether, these comparisons demonstrate the potential of our experimental pipeline to characterize the most insoluble ECM components not only from soft tissues as previously shown but also from mineralized tissues.

Conclusion

We propose that the comprehensive list of core ECM and ECM-associated proteins of young and old human root dentin provided here, constitutes the first step towards the discovery of potential biomarkers of the aging tooth. While our initial in-depth study of the dentin

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

matrisome of two age brackets revealed age-related changes to dentin ECM protein solubility and composition, investigations are now necessary to unveil age-dependent mechanisms leading to these changes. For example, changes in ECM protein solubility can result from alterations in the level of cross-linking or degradation of these proteins, while changes in the composition can result from modifications in gene expression or protein secretion. Further investigations will also be needed to address the functional consequences of the changes in ECM solubility and composition during aging. If successful, they may lead to the development of novel preventative strategies to maintain a healthy dentition and improve the health of our aging society.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

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 May 28, 2020. 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

(1) United Nations, Department of Economic and Social Affairs. *World Population Prospects 2019*; 2019.

(2) Stock, C.; Jürges, H.; Shen, J.; Bozorgmehr, K.; Listl, S. A Comparison of Tooth Retention and Replacement across 15 Countries in the Over-50s. *Community Dent Oral Epidemiol* **2015**, n/a-n/a. https://doi.org/10.1111/cdoe.12209.

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

(4) McNally, M. E.; Matthews, D. C.; Clovis, J. B.; Brillant, M.; Filiaggi, M. J. The Oral Health of Ageing Baby Boomers: A Comparison of Adults Aged 45-64 and Those 65 Years and Older. *Gerodontology* **2014**, *31* (2), 123–135. https://doi.org/10.1111/ger.12022.

(5) Butler, W. T.; Brunn, J. C.; Qin, C. Dentin Extracellular Matrix (ECM) Proteins: Comparison to Bone ECM and Contribution to Dynamics of Dentinogenesis. *Connective Tissue Research* **2003**, *44* (1), 171–178. https://doi.org/10.1080/03008200390152287.

(6) Alford, A. I.; Kozloff, K. M.; Hankenson, K. D. Extracellular Matrix Networks in Bone Remodeling. *Int. J. Biochem. Cell Biol.* **2015**, *65*, 20–31. https://doi.org/10.1016/j.biocel.2015.05.008.

(7) Blair, H. C.; Larrouture, Q. C.; Li, Y.; Lin, H.; Beer-Stoltz, D.; Liu, L.; Tuan, R. S.; Robinson, L. J.; Schlesinger, P. H.; Nelson, D. J. Osteoblast Differentiation and Bone Matrix Formation *In Vivo* and *In Vitro*. *Tissue Engineering Part B: Reviews* **2017**, *23* (3), 268–280. https://doi.org/10.1089/ten.teb.2016.0454.

(8) Szulc, P. Bone Turnover: Biology and Assessment Tools. *Best Pract. Res. Clin. Endocrinol. Metab.* **2018**, *32* (5), 725–738. https://doi.org/10.1016/j.beem.2018.05.003.

(9) Goldberg, M.; Kulkarni, A. B.; Young, M.; Boskey, A. Dentin: Structure, Composition and Mineralization. *Front Biosci (Elite Ed)* **2011**, *3*, 711–735. https://doi.org/10.2741/e281.

(10) Avery, J. K. In *Oral development and histology*; Steele, P. F., Avery, N., Eds.; Thieme: Stuttgart ; New York, 2002; pp 72–108; 172–190.

(11) Bertassoni, L. E.; Orgel, J. P. R.; Antipova, O.; Swain, M. V. The Dentin Organic Matrix – Limitations of Restorative Dentistry Hidden on the Nanometer Scale. *Acta Biomaterialia* **2012**, *8* (7), 2419–2433. https://doi.org/10.1016/j.actbio.2012.02.022.

(12) Marshall, G. W. Dentin: Microstructure and Characterization. *Quintessence Int* **1993**, *24* (9), 606–617.

(13) Ricard-Blum, S. The Collagen Family. *Cold Spring Harbor Perspectives in Biology* **2011**, *3* (1), a004978–a004978. https://doi.org/10.1101/cshperspect.a004978.

(14) 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. Proteomics Analysis of Human Dentin Reveals Distinct Protein Expression Profiles. *J. Proteome Res.* **2009**, *8* (3), 1338–1346. https://doi.org/10.1021/pr801065s.

(15) Jágr, M.; Eckhardt, A.; Pataridis, S.; Mikšík, I. Comprehensive Proteomic Analysis of Human Dentin. *Eur J Oral Sci* **2012**, *120* (4), 259–268. https://doi.org/10.1111/j.1600-0722.2012.00977.x.

(16) Linde, A. Dentin Matrix Proteins: Composition and Possible Functions in Calcification. *Anat. Rec.* **1989**, *224* (2), 154–166. https://doi.org/10.1002/ar.1092240206.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

(17) George, A.; Sabsay, B.; Simonian, P. A.; Veis, A. Characterization of a Novel Dentin Matrix Acidic Phosphoprotein. Implications for Induction of Biomineralization. *J. Biol. Chem.* **1993**, *268* (17), 12624–12630.
(18) Boskey, A. L. The Role of Extracellular Matrix Components in Dentin Mineralization. *Crit. Rev. Oral Biol. Med.* **1991**, *2* (3), 369–387. https://doi.org/10.1177/10454411910020030501.
(19) Butler, W. T. Dentin Matrix Proteins. *European Journal of Oral Sciences* **1998**, *106* (S1), 204–210. https://doi.org/10.1111/j.1600-0722.1998.tb02177.x.
(20) Ravindran, S.; George, A. Multifunctional ECM Proteins in Bone and Teeth. *Experimental Cell Research* **2014**, *325* (2), 148–154. https://doi.org/10.1016/j.yexcr.2014.01.018.
(21) Embery, G.; Hall, R.; Waddington, R.; Septier, D.; Goldberg, M. Proteoglycans in Dentinogenesis. *Crit. Rev. Oral Biol. Med.* **2001**, *12* (4), 331–349. https://doi.org/10.1177/10454411010120040401.
(22) Goldberg, M.; Takagi, M. Dentine Proteoglycans: Composition, Ultrastructure and Functions. *Histochem J* **1993**, *25* (11), 781–806. https://doi.org/10.1007/BF02388111.
(23) Iozzo, R. V.; Schaefer, L. Proteoglycan Form and Function: A Comprehensive Nomenclature of Proteoglycans. *Matrix Biology* **2015**, *42*, 11–55. https://doi.org/10.1016/j.matbio.2015.02.003.
(24) Bertassoni, L. E.; Swain, M. V. The Contribution of Proteoglycans to the Mechanical Behavior of Mineralized Tissues. *J Mech Behav Biomed Mater* **2014**, *38*, 91–104. https://doi.org/10.1016/j.jmbbm.2014.06.008.
(25) de Mattos Pimenta Vidal, C.; Leme-Kraus, A. A.; Rahman, M.; Farina, A. P.; Bedran-Russo, A. K. Role of Proteoglycans on the Biochemical and Biomechanical Properties of Dentin Organic Matrix. *Arch. Oral Biol.* **2017**, *82*, 203–208. https://doi.org/10.1016/j.archoralbio.2017.06.020.
(26) Tjäderhane, L.; Buzalaf, M. A. R.; Carrilho, M.; Chaussain, C. Matrix Metalloproteinases and Other Matrix Proteinases in Relation to Cariology: The Era of "Dentin Degradomics." *Caries Res* **2015**, *49* (3), 193–208. https://doi.org/10.1159/000363582.
(27) Kinney, J. H.; Nalla, R. K.; Pople, J. A.; Breunig, T. M.; Ritchie, R. O. Age-Related Transparent Root Dentin: Mineral Concentration, Crystallite Size, and Mechanical Properties. *Biomaterials* **2005**, *26* (16), 3363–3376. https://doi.org/10.1016/j.biomaterials.2004.09.004.
(28) Kabartai, F.; Hoffmann, T.; Hannig, C. The Physiologic Sclerotic Dentin: A Literature-Based Hypothesis. *Medical Hypotheses* **2015**, *85* (6), 887–890. https://doi.org/10.1016/j.mehy.2015.09.016.
(29) Orofacial Development & Regeneration, Institute of Oral Biology, Centre for Dental Medicine, Faculty of Medicine, University of Zurich, Plattenstrasse 11, 8032 Zurich, Switzerland.; Iezzi, I.; Pagella, P.; Mattioli-Belmonte, M.; Mitsiadis, T. The Effects of Ageing on Dental Pulp Stem Cells, the Tooth Longevity Elixir. *eCM* **2019**, *37*, 175–185. https://doi.org/10.22203/eCM.v037a11.
(30) Nazari, A.; Bajaj, D.; Zhang, D.; Romberg, E.; Arola, D. Aging and the Reduction in Fracture Toughness of Human Dentin. *Journal of the Mechanical Behavior of Biomedical Materials* **2009**, *2* (5), 550–559. https://doi.org/10.1016/j.jmbbm.2009.01.008.
(31) Arola, D.; Reprogel, R. K. Effects of Aging on the Mechanical Behavior of Human Dentin. *Biomaterials* **2005**, *26* (18), 4051–4061. https://doi.org/10.1016/j.biomaterials.2004.10.029.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

(32) Ryou, H.; Romberg, E.; Pashley, D. H.; Tay, F. R.; Arola, D. Importance of Age on the Dynamic Mechanical Behavior of Intertubular and Peritubular Dentin. *Journal of the Mechanical Behavior of Biomedical Materials* **2015**, *42*, 229–242. https://doi.org/10.1016/j.jmbbm.2014.11.021.

(33) Martin-De Las Heras, S.; Valenzuela, A.; Overall, C. M. The Matrix Metalloproteinase Gelatinase A in Human Dentine. *Archives of Oral Biology* **2000**, *45* (9), 757–765. https://doi.org/10.1016/S0003-9969(00)00052-2.

(34) Schiller, H. B.; Fernandez, I. E.; Burgstaller, G.; Schaab, C.; Scheltema, R. A.; Schwarzmayr, T.; Strom, T. M.; Eickelberg, O.; Mann, M. Time- and Compartment-Resolved Proteome Profiling of the Extracellular Niche in Lung Injury and Repair. *Molecular Systems Biology* **2015**, *11* (7), 819. https://doi.org/10.15252/msb.20156123.

(35) Randles, M.; Lennon, R. Applying Proteomics to Investigate Extracellular Matrix in Health and Disease. *Curr Top Membr* **2015**, *76*, 171–196. https://doi.org/10.1016/bs.ctm.2015.06.001.

(36) Taha, I. N.; Naba, A. Exploring the Extracellular Matrix in Health and Disease Using Proteomics. *Essays In Biochemistry* **2019**, EBC20190001. https://doi.org/10.1042/EBC20190001.

(37) Naba, A.; Clauser, K. R.; Ding, H.; Whittaker, C. A.; Carr, S. A.; Hynes, R. O. The Extracellular Matrix: Tools and Insights for the "Omics" Era. *Matrix Biology* **2016**, *49*, 10–24. https://doi.org/10.1016/j.matbio.2015.06.003.

(38) Socovich, A. M.; Naba, A. The Cancer Matrisome: From Comprehensive Characterization to Biomarker Discovery. *Seminars in Cell & Developmental Biology* **2019**, *89*, 157–166. https://doi.org/10.1016/j.semcdb.2018.06.005.

(39) Grilo, G. A.; Shaver, P. R.; Stoffel, H. J.; Morrow, C. A.; Johnson, O. T.; Iyer, R. P.; de Castro Brás, L. E. Age- and Sex-Dependent Differences in Extracellular Matrix Metabolism Associate with Cardiac Functional and Structural Changes. *Journal of Molecular and Cellular Cardiology* **2020**, *139*, 62–74. https://doi.org/10.1016/j.yjmcc.2020.01.005.

(40) 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. An Atlas of the Aging Lung Mapped by Single Cell Transcriptomics and Deep Tissue Proteomics. *Nat Commun* **2019**, *10* (1), 963. https://doi.org/10.1038/s41467-019-08831-9.

(41) Naba, A.; Clauser, K. R.; Hoersch, S.; Liu, H.; Carr, S. A.; Hynes, R. O. The Matrisome: In Silico Definition and in Vivo Characterization by Proteomics of Normal and Tumor Extracellular Matrices. *Mol Cell Proteomics* **2012**, *11* (4), M111.014647. https://doi.org/10.1074/mcp.M111.014647.

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

(43) 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.; Ishihama, Y.; Vizcaíno, J. A. The ProteomeXchange Consortium in 2020: Enabling 'Big Data' Approaches in Proteomics. *Nucleic Acids Res* **2020**, *48* (D1), D1145–D1152. https://doi.org/10.1093/nar/gkz984.

(44) 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.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaíno, J. A. The PRIDE Database and

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

Related Tools and Resources in 2019: Improving Support for Quantification Data. *Nucleic Acids Res* **2019**, *47* (D1), D442–D450. https://doi.org/10.1093/nar/gky1106.

(45) Peffers, M. J.; Thorpe, C. T.; Collins, J. A.; Eong, R.; Wei, T. K. J.; Screen, H. R. C.; Clegg, P. D. Proteomic Analysis Reveals Age-Related Changes in Tendon Matrix Composition, with Age- and Injury-Specific Matrix Fragmentation. *J. Biol. Chem.* **2014**, *289* (37), 25867–25878. https://doi.org/10.1074/jbc.M114.566554.

(46) Gelse, K. Collagens—Structure, Function, and Biosynthesis. *Advanced Drug Delivery Reviews* **2003**, *55* (12), 1531–1546. https://doi.org/10.1016/j.addr.2003.08.002.

(47) Ritchie, H. The Functional Significance of Dentin Sialoprotein-Phosphophoryn and Dentin Sialoprotein. *Int J Oral Sci* **2018**, *10* (4), 31. https://doi.org/10.1038/s41368-018-0035-9.

(48) Kawasaki, K.; Suzuki, T.; Weiss, K. M. Genetic Basis for the Evolution of Vertebrate Mineralized Tissue. *Proceedings of the National Academy of Sciences* **2004**, *101* (31), 11356–11361. https://doi.org/10.1073/pnas.0404279101.

(49) Schaefer, L.; Iozzo, R. V. Biological Functions of the Small Leucine-Rich Proteoglycans: From Genetics to Signal Transduction. *J. Biol. Chem.* **2008**, *283* (31), 21305–21309. https://doi.org/10.1074/jbc.R800020200.

(50) Gubbiotti, M. A.; Neill, T.; Iozzo, R. V. A Current View of Perlecan in Physiology and Pathology: A Mosaic of Functions. *Matrix Biology* **2017**, *57–58*, 285–298. https://doi.org/10.1016/j.matbio.2016.09.003.

(51) Edwards, D.; Handsley, M.; Pennington, C. The ADAM Metalloproteinases. *Molecular Aspects of Medicine* **2008**, *29* (5), 258–289. https://doi.org/10.1016/j.mam.2008.08.001.

(52) Yazdani, U.; Terman, J. R. The Semaphorins. *Genome Biology* **2006**, *7* (3), 211. https://doi.org/10.1186/gb-2006-7-3-211.

(53) Zalzal, S.; Smith, C.; Nanci, A. Ameloblastin and Amelogenin Share a Common Secretory Pathway and Are Co-Secreted during Enamel Formation. *Matrix Biology* **2008**, *27* (4), 352–359. https://doi.org/10.1016/j.matbio.2008.01.003.

(54) Atsawasuwan, P.; Lu, X.; Ito, Y.; Chen, Y.; Gopinathan, G.; Evans, C. A.; Kulkarni, A. B.; Gibson, C. W.; Luan, X.; Diekwisch, T. G. H. Expression and Function of Enamel-Related Gene Products in Calvarial Development. *J Dent Res* **2013**, *92* (7), 622–628. https://doi.org/10.1177/0022034513487906.

(55) Jacques, J.; Hotton, D.; Asselin, A.; De la Dure-Molla, M.; Coudert, A. E.; Isaac, J.; Berdal, A. Ameloblastin as a Putative Marker of Specific Bone Compartments. *Connective Tissue Research* **2014**, *55* (sup1), 117–120. https://doi.org/10.3109/03008207.2014.923849.

(56) Carrino, D. A.; Calabro, A.; Darr, A. B.; Dours-Zimmermann, M. T.; Sandy, J. D.; Zimmermann, D. R.; Sorrell, J. M.; Hascall, V. C.; Caplan, A. I. Age-Related Differences in Human Skin Proteoglycans. *Glycobiology* **2011**, *21* (2), 257–268. https://doi.org/10.1093/glycob/cwq162.

(57) Bonnans, C.; Chou, J.; Werb, Z. Remodelling the Extracellular Matrix in Development and Disease. *Nat Rev Mol Cell Biol* **2014**, *15* (12), 786–801. https://doi.org/10.1038/nrm3904.

(58) Copes, N.; Edwards, C.; Chaput, D.; Saifee, M.; Barjuca, I.; Nelson, D.; Paraggio, A.; Saad, P.; Lipps, D.; Stevens, S. M.; Bradshaw, P. C. Metabolome and Proteome Changes with Aging in Caenorhabditis Elegans. *Exp. Gerontol.* **2015**, *72*, 67–84. https://doi.org/10.1016/j.exger.2015.09.013.

(59) Chaleckis, R.; Murakami, I.; Takada, J.; Kondoh, H.; Yanagida, M. Individual Variability in Human Blood Metabolites Identifies Age-Related Differences. *Proc Natl Acad Sci USA* **2016**, *113* (16), 4252–4259. https://doi.org/10.1073/pnas.1603023113.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.27.116624; this version posted May 28, 2020. 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.

(60) Wang, C.; Brisson, B. K.; Terajima, M.; Li, Q.; Hoxha, K.; Han, B.; Goldberg, A. M.; Sherry Liu, X.; Marcolongo, M. S.; Enomoto-Iwamoto, M.; Yamauchi, M.; Volk, S. W.; Han, L. Type III Collagen Is a Key Regulator of the Collagen Fibrillar Structure and Biomechanics of Articular Cartilage and Meniscus. *Matrix Biology* **2020**, *85–86*, 47–67. https://doi.org/10.1016/j.matbio.2019.10.001.

(61) Malfait, F. Vascular Aspects of the Ehlers-Danlos Syndromes. *Matrix Biology* **2018**, *71–72*, 380–395. https://doi.org/10.1016/j.matbio.2018.04.013.

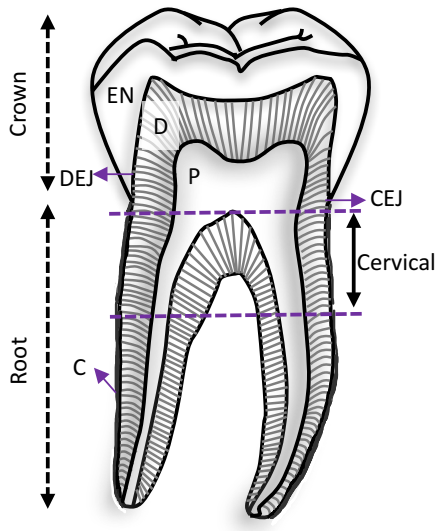
(62) Liu, X.; Wu, H.; Byrne, M.; Krane, S.; Jaenisch, R. Type III Collagen Is Crucial for Collagen I Fibrillogenesis and for Normal Cardiovascular Development. *Proceedings of the National Academy of Sciences* **1997**, *94* (5), 1852–1856. https://doi.org/10.1073/pnas.94.5.1852.

(63) Naba, A.; Clauser, K. R.; Ding, H.; Whittaker, C. A.; Carr, S. A.; Hynes, R. O. The Extracellular Matrix: Tools and Insights for the "Omics" Era. *Matrix Biology* **2016**, *49*, 10–24. https://doi.org/10.1016/j.matbio.2015.06.003.

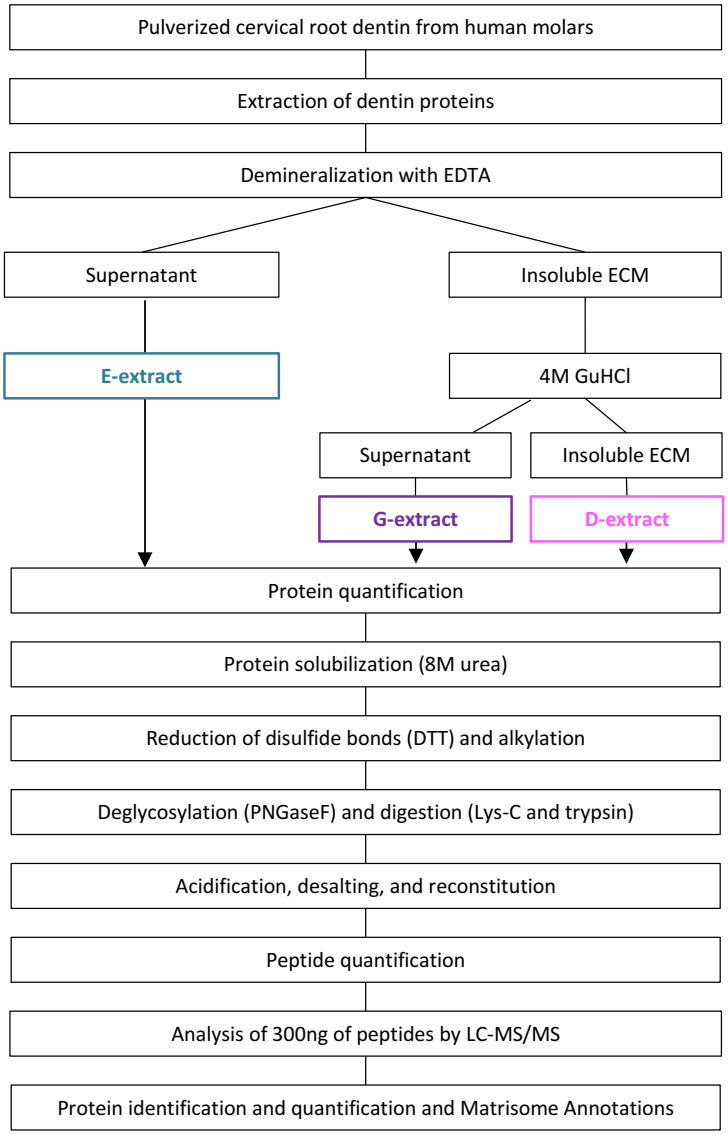
(64) Naba, A.; Pearce, O. M. T.; Del Rosario, A.; Ma, D.; Ding, H.; Rajeeve, V.; Cutillas, P. R.; Balkwill, F. R.; Hynes, R. O. Characterization of the Extracellular Matrix of Normal and Diseased Tissues Using Proteomics. *J. Proteome Res.* **2017**, *16* (8), 3083–3091. https://doi.org/10.1021/acs.jproteome.7b00191.

Figure 1.

A.



B.



C.

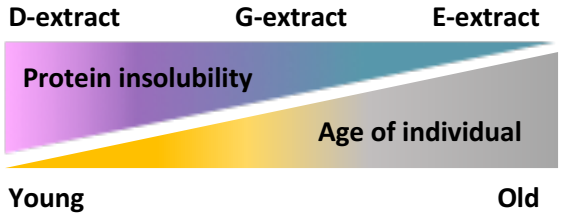


Figure 1. ECM Enrichment pipeline for the cervical root dentin. (**A**) Schematic of a sound human molar (EN: enamel, D: dentin, C: cementum, DEJ: dentin-enamel junction, CEJ: cement-enamel junction, and P: pulp) and the type of dentin (root) and area (cervical) used for protein extraction. (**B**) Experimental pipeline describing sample preparation, protein extraction, digestion, and mass spectrometry analysis. (**C**) Color code used throughout the figures to identify the different protein extracts and the age of the individual.

Figure 2.

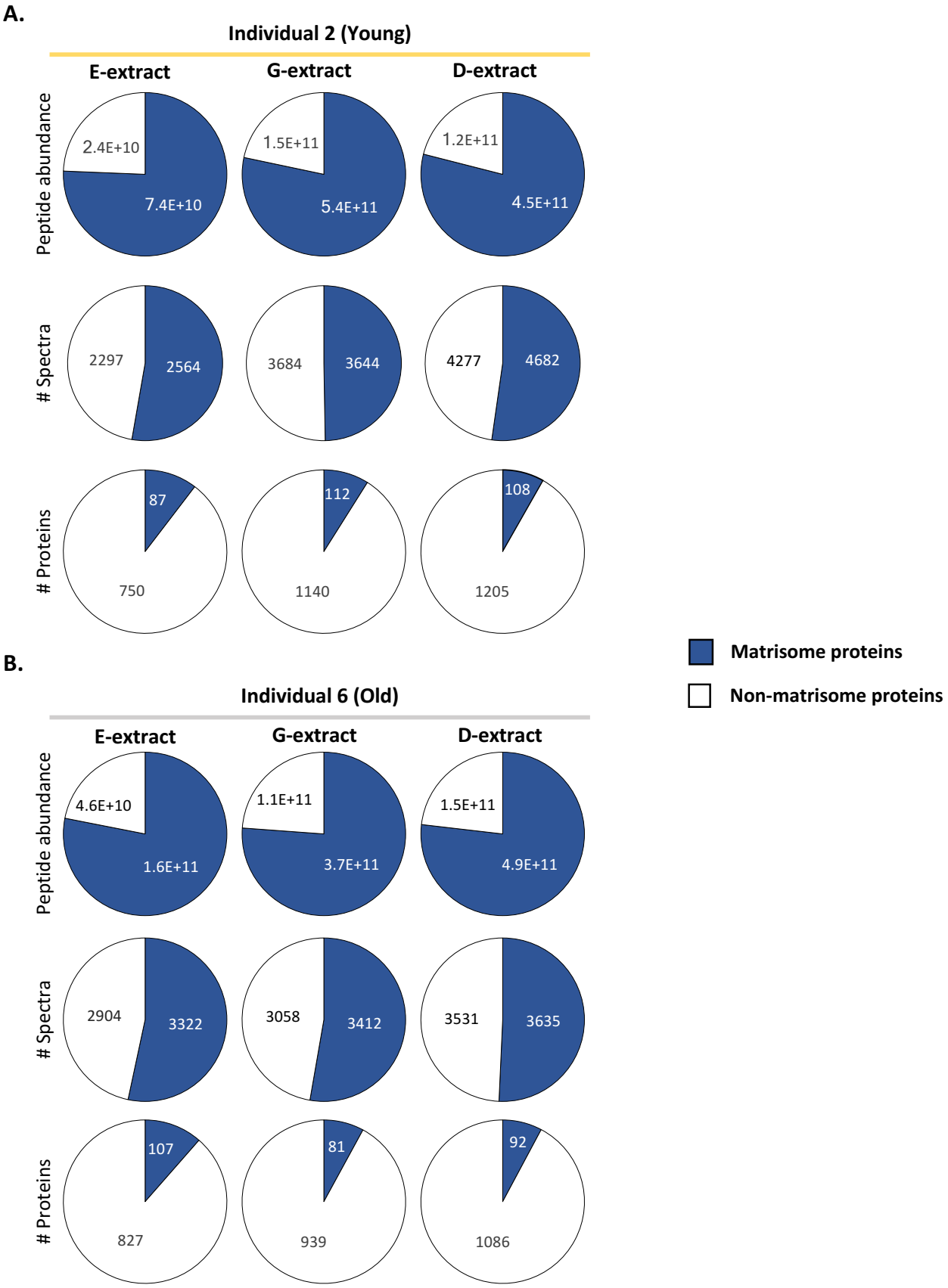


Figure 2. Characterization of ECM proteins abundance in extracts of different solubility in young and old dentin samples. Pie charts show peptide abundance (upper panels), number of total spectra (middle panels), and number of proteins (lower panels) across all extracts (E, G, and D) for representative young (**A**) and old (**B**) samples (related to Supplementary Table 1A).

Figure 3.

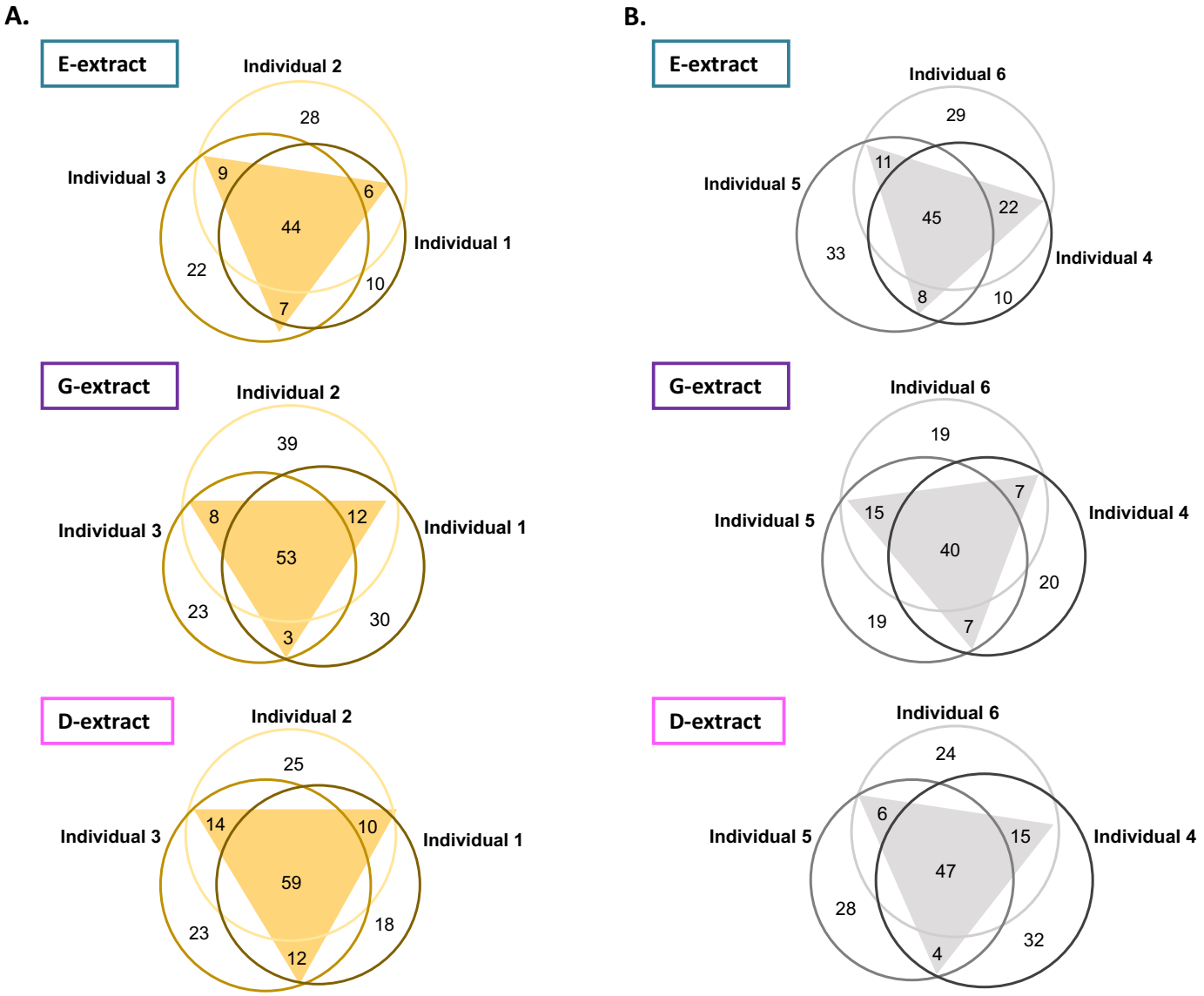


Figure 3. Inter-individual comparison of ECM proteins detected in extracts of differing solubility and definition of the young and old dentin matrisomes. Venn diagrams display the inter-individual comparisons of young (**A**) and old (**B**) ECM proteins detected in E- (upper panels), G- (middle panels), and D- (lower panels) extracts. The triangles highlight the ensemble of proteins used to define the matrisomes each extract and consist of ECM proteins identified in at least two individuals.

Figure 4.

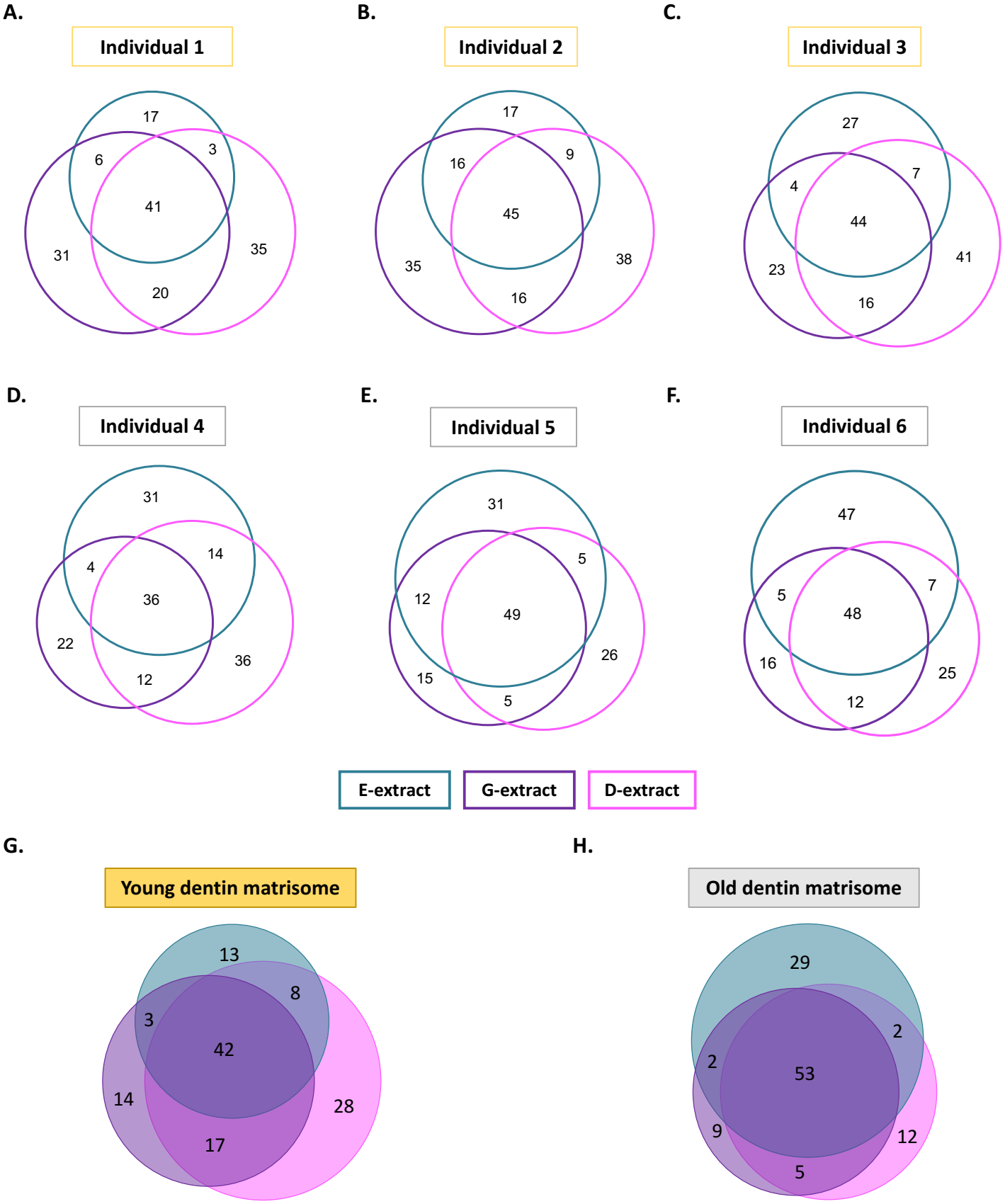
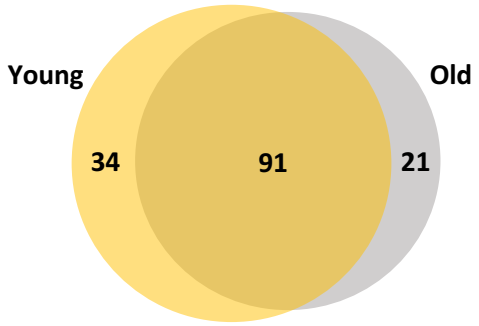


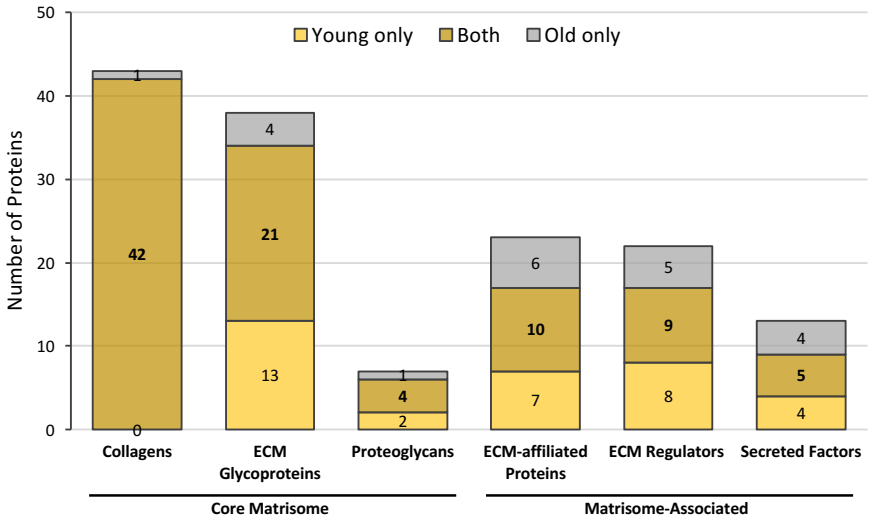
Figure 4. Comparison of young and old dentin ECM proteins detected in extracts of different solubility and definition of the dentin matrisome. Venn diagrams display the number of matrisome proteins detected E-, G-, and D-extracts in young (**A - C**) and old (**D - F)** samples (related to Supplementary Tables 1E to 1L). Proteins included in the analysis were detected with at least two peptides. (**G - H)** The young and old dentin matrisomes are defined as the ensemble of proteins detected in at least one of the E-, G-, or D- extract and at least in 2 of 3 individuals (related to Supplementary Tables 1H and 1M).

Figure 5.

A.



B.



C.

ECM proteins only detected in young individuals	
Core Matrisome	**Matrisome-Associated**
ECM Glycoproteins	***ECM-affiliated Proteins***
AMBN	C1QTNF2
AMELY	CLEC4F
CRELD1	FREM3
DPT	GPC3
EMID1	MUC5B
FNDC1	SEMA4G
FNDC8	SEMA5B
LAMA2	***ECM Regulators***
LAMA3	ADAM2
OTOG	ADAMTS12
POSTN	ADAMTS2
RELN	ADAMTSL3
SSPO	CTSD
Proteoglycans	LOXL2
DCN	SERPINA10
NCAN	TLL2
	Secreted Factors
	FGF11
	FLG
	IL3
	INS-IGF2

D.

ECM proteins only detected in old individuals	
Core Matrisome	**Matrisome-Associated**
Collagens	***ECM-affiliated Proteins***
COL6A1	C1QB
ECM Glycoproteins	C1QL2
EMILIN3	C1QL3
HMCN1	GPC6
LAMB3	MUC19
TNR	MUC5AC
Proteoglycans	***ECM Regulators***
PRG4	ADAM28
	ADAM8
	LOX
	SERPINB5
	SERPINC1
	Secreted Factors
	IGF2
	IL25
	NRG1
	PTN

Figure 5. Comparison of the young and old dentin matrisomes. (**A**) Venn diagram illustrates the number of matrisome proteins detected in either or both age groups (related to Supplementary Tables 1D, 1I, and 1N). (**B**) Bar chart shows the distribution of proteins detected in young (yellow) and old (grey) dentin matrisomes across the different matrisome categories (collagens, ECM glycoproteins, proteoglycans, ECM-affiliated proteins, ECM regulators, and secreted factors). (**C**) List of 34 ECM proteins uniquely identified in young individuals. (**D**) List of 21 proteins uniquely identified in old individuals. Proteins are classified in matrisome divisions and categories (related to Supplementary Tables 1D, 1I, and 1N).

Table 1: Matrisome of young and old human dentin

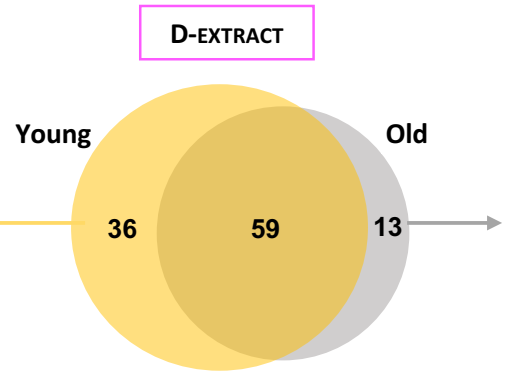
		Young			Old		
	Gene Name	E-extract	G-extract	D-extract	E-extract	G-extract	D-extract
Collagens	COL10A1						
	COL11A1						
	COL11A2						
	COL12A1						
	COL13A1						
	COL15A1						
	COL16A1						
	COL19A1						
	COL1A1						
	COL1A2						
	COL21A1						
	COL22A1						
	COL23A1						
	COL24A1						
	COL25A1						
	COL27A1						
	COL28A1						
	COL2A1						
	COL3A1						
	COL4A1						
	COL4A2						
	COL4A3						
	COL4A4						
	COL4A5						
	COL4A6						
	COL5A1						
	COL5A2						
	COL5A3						
	COL6A5						
	COL7A1						
	COL8A1						
	COL8A2						
	COL9A1						
	COL9A2						
	COL6A2						
	COL14A1						
	COL17A1						
	COL18A1						
	COL26A1						
	COL9A3						
	COL6A3						
	COL6A6						
	COL6A1						
Proteoglycans	BGN						
	OMD						
	HSPG2						
	LUM						
	DCN						
	NCAN						
	PRG4						

		Young			Old		
	Gene Name	E-extract	G-extract	D-extract	E-extract	G-extract	D-extract
ECM Glycoproteins	EMILIN2						
	VTN						
	MGP						
	PCOLCE						
	OTOL1						
	DSPP						
	FRAS1						
	IGFBP5						
	SPARC						
	SPP1						
	CTHRC1						
	COLQ						
	GLDN						
	LAMA5						
	MXRA5						
	ECM1						
	FN1						
	HMCN2						
	FBLN2						
	CRELD1						
	DPT						
	FNDC1						
	LAMA3						
	RELN						
	SSPO						
	TNXB						
	IGSF10						
	AMBN						
	AMELY						
	EMID1						
	FNDC8						
	LAMA2						
	OTOG						
	POSTN						
	LAMB3						
	HMCN1						
	TNR						
	EMILIN3						
Secreted Factors	CCBE1						
	TCHHL1						
	WNT5B						
	C1QTNF9B						
	INS-IGF2						
	EDA						
	FGF11						
	FLG						
	IL3						
	IGF2						
	NRG1						
	PTN						
	IL25						

		Young			Old		
	Gene Name	E-extract	G-extract	D-extract	E-extract	G-extract	D-extract
ECM-affiliated proteins	MUC16						
	LMAN1L						
	SEMA6D						
	CLEC3B						
	COLEC12						
	SFTPD						
	MBL2						
	C1QTNF9						
	FCN2						
	MUC5B						
	LGALS9						
	C1QTNF2						
	CLEC4F						
	FREM3						
	GPC3						
	SEMA4G						
	SEMA5B						
	GPC6						
	MUC5AC						
	C1QL3						
	C1QB						
	C1QL2						
	MUC19						
ECM Regulators	F2						
	SERPINF1						
	MMP20						
	KNG1						
	ADAM29						
	F10						
	F9						
	SERPINA1						
	ADAMTS12						
	ADAMTSL5						
	LOXL2						
	ADAM2						
	ADAMTS2						
	ADAMTSL3						
	CTSD						
	SERPINA10						
	TLL2						
	LOX						
	SERPINC1						
	ADAM28						
	ADAM8						
	SERPINB5						

The table contains the matrisome of E-, G-, and D- extracts, each consisting of ECM proteins identified in at least two individuals. Proteins are grouped by the following matrisome categories: collagens, ECM glycoproteins, proteoglycans, ECM-affiliated proteins, ECM regulators and secreted factors. Cells highlighted in yellow and gray indicate the ECM proteins detected in young and old individuals, respectively.

Figure 6.



ECM proteins only detected in D-extract of young individuals

Core Matrisome	Matrisome-Associated
Collagens	***ECM-affiliated Proteins***
COL6A6	C1QTNF2
ECM Glycoproteins	CLEC4F
AMBN	FREM3
AMELY	GPC3
ECM1	LGALS9
EMID1	LMAN1L
FN1	SEMA4G
FNDC8	SEMA5B
LAMA2	***ECM Regulators***
OTOG	ADAM2
OTOL1	ADAMTS2
POSTN	ADAMTSL3
VTN	ADAMTSL5
Proteoglycans	CTSD
NCAN	MMP20
OMD	SERPINA10
	SERPINF1
	TLL2
	Secreted Factors
	EDA
	FGF11
	FLG
	IL3
	TCHHL1

ECM proteins only detected in D-extract of old individuals

Core Matrisome	Matrisome-Associated
Collagens	***ECM-affiliated Proteins***
COL6A3	C1QB
ECM Glycoproteins	C1QL2
EMILIN3	C1QTNF9
FBLN2	MUC19
Proteoglycans	***ECM Regulators***
PRG4	ADAM28
	ADAM8
	SERPINB5
	Secreted Factors
	C1QTNF9B
	IL25

Figure 6. Analysis of the ECM proteins identified in D-extracts of young and old individuals. Venn diagram represents the overlap between ECM proteins found in young and old D-extracts. Tables list ECM proteins detected exclusively in young (yellow, left panel) or old (gray, right panel) D-extracts. The complete young and old human dentin matrisomes are presented Table 1 and Supplementary Table 1N. Proteins are grouped into matrisome divisions and categories.

Figure 7.

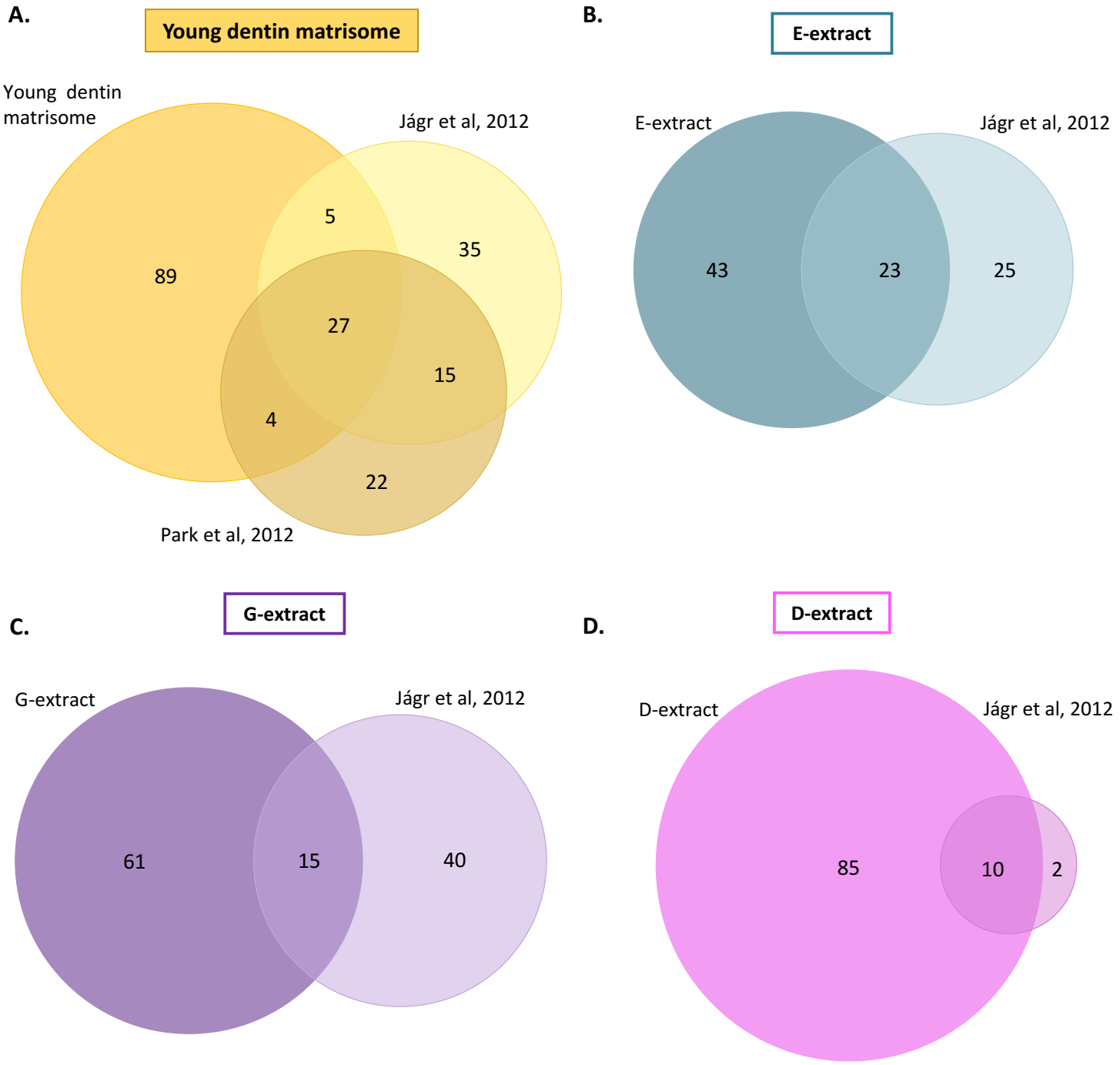


Figure 7. Comparisons of young matrisome with those of published studies, with matrisome annotations. The Venn diagrams illustrate the number of shared and distinct proteins between the lists defined as the young dentin matrisome in this study and the re-annotated proteomes of human dentin published by Jágr et al, and Park et al **(A)**. With differing methods of extraction used, the subsequent breakdown is further grouped and illustrated by: E-extracts **(B)**, G-extracts **(C)**, and D-extracts **(D)**. The complete matrisome of each study is presented in Supplementary Table 2, where the proteins are grouped into matrisome divisions and categories.

Table 2: Matrisomes of young human dentin

All extracts				
	Gene Name	Young Dentin Matrisome	Jágr et al, 2012	Park et al, 2009
ECM Glycoproteins	AEBP1			■
	AMBN	■		
	AMELY	■		
	BGLAP		■	
	COLQ	■		
	CRELD1	■		
	CTHRC1	■		
	DMBT1		■	
	DMP1		■	
	DPT	■	■	■
	DSPP	■	■	■
	ECM1	■		
	ECM2		■	■
	EMID1	■		
	EMILIN2	■		
	FBLN2	■		
	FGB		■	
	FGG		■	
	FN1	■		■
	FNDC1	■		
	FNDC8	■		
	FRAS1	■		
	GLDN	■		
	HMCN2	■		
	IBSP			■
	IGFBP1		■	
	IGFBP3		■	
	IGFBP5	■	■	■
	IGSF10	■		
	LAMA2	■		
	LAMA3	■		
	LAMA5	■		
	LRG1		■	
	LTBP3		■	
	MEPE		■	
	MFGE8			■
	MGP	■	■	■
	MXRA5	■		
	OTOG	■		
	OTOL1	■		
	PCOLCE	■	■	■
	POSTN	■	■	■
	RELN	■		
	SPARC	■	■	■
	SPP1	■	■	■
	SRPX			■
	SSPO	■		
	TGFBI		■	■
	THBS1		■	■
	TNC		■	■
	TNN			■
	TNXB	■		
	VTN	■	■	■
Proteoglycans	ASPN		■	■
	BGN	■	■	■
	CHAD		■	■
	DCN	■	■	■
	FMOD		■	■
	HAPLN1			■
	HSPG2	■	■	■
	LUM	■	■	■
	NCAN	■		
	OGN		■	■
	OMD	■	■	■
	PRELP			■
	SPOCK3		■	
	VCAN			■

	Gene Name	Young Dentin Matrisome	Jágr et al, 2012	Park et al, 2009
Collagens	COL10A1	■		
	COL11A1	■		
	COL11A2	■	■	■
	COL12A1	■	■	■
	COL13A1	■		
	COL14A1	■		
	COL15A1	■		
	COL16A1	■		
	COL17A1	■		
	COL18A1	■		
	COL19A1	■		
	COL1A1	■	■	■
	COL1A2	■	■	■
	COL21A1	■		
	COL22A1	■		
	COL23A1	■		
	COL24A1	■		
	COL25A1	■		
	COL26A1	■		
	COL27A1	■		
	COL28A1	■		
	COL2A1	■	■	
	COL3A1	■	■	
	COL4A1	■		
	COL4A2	■		
	COL4A3	■		
	COL4A4	■		■
	COL4A5	■		
	COL4A6	■		
	COL5A1	■	■	■
	COL5A2	■	■	
	COL5A3	■		
	COL6A1			■
	COL6A2	■		■
	COL6A3	■		■
	COL6A5	■		
	COL6A6	■		
	COL7A1	■		
	COL8A1	■		
	COL8A2	■		
	COL9A1	■		
	COL9A2	■		
	COL9A3	■		
Secreted Factors	C1QTNF9B	■		
	CCBE1	■		
	EDA	■		
	FGF11	■		
	FLG	■		
	FLG2			■
	HRNR		■	
	IL3	■		
	INS-IGF2	■		
	MDK		■	
	PTN		■	
	S100A6		■	
	S100A7		■	
	S100A8		■	
	S100A9		■	
	TCHHL1	■		
	TGFB1		■	■
	WNT5B	■		

	Gene Name	Young Dentin Matrisome	Jágr et al, 2012	Park et al, 2009
ECM-affiliated proteins	ANXA1		■	
	ANXA11			■
	ANXA2		■	■
	ANXA4			■
	ANXA5		■	■
	ANXA6			■
	ANXA7			■
	C1QTNF2	■		
	C1QTNF9	■		
	CLEC11A		■	■
	CLEC3B	■	■	■
	CLEC4F	■		
	COLEC12	■		
	FCN2	■		
	FREM3	■		
	GPC3	■		
	HPX		■	
	LGALS1		■	
	LGALS3			■
	LGALS7		■	
	LGALS9	■		
	LMAN1L	■		
	MBL2	■		
	MUC16	■		
	MUC5B	■		
	MUC6		■	
	SEMA4G	■		
	SEMA5B	■		
	SEMA6D	■		
	SFTPD	■		
ECM Regulators	ADAM2	■		
	ADAM29	■		
	ADAMTS12	■		
	ADAMTS2	■		
	ADAMTSL3	■		
	ADAMTSL5	■		
	AGT		■	
	AMBP		■	
	CTSD	■	■	
	F10	■	■	■
	F2	■	■	■
	F7		■	■
	F9	■	■	■
	HRG		■	
	ITIH2			■
	ITIH3		■	
	ITIH4		■	
	KNG1	■	■	■
	LOX			■
	LOXL2	■		
	MMP2			■
	MMP20	■	■	■
	PLAT			■
	PLG			■
	SERPINA1	■	■	■
	SERPINA10	■	■	
	SERPINA3		■	■
	SERPINA6		■	
	SERPINB3		■	
	SERPINB4		■	
	SERPINC1		■	■
	SERPINF1	■	■	■
	SERPINF2		■	
	SERPING1		■	
	SERPINH1			■
	SLPI			■
	TGM3		■	
	TIMP1		■	
	TLL2	■		

The table contains the matrisome of the young dentin, and re-annotated studies done by Jágr et al (2012), and Park et al (2009). Proteins are grouped by the following matrisome categories: collagens, ECM glycoproteins, proteoglycans, ECM-affiliated proteins, ECM regulators and secreted factors. Cells highlighted in yellow and brown indicate the ECM proteins detected in this study and previous, respectively.