Review

The extracellular matrix: Tools and insights for the "omics" era

Alexandra Naba[a], Karl R. Clauser[c], Huiming Ding[d], Charles A. Whittaker[d], Steven A. Carr[c] and Richard O. Hynes[a,b,]

a - David H. Koch Institute for Integrative Cancer Research, *Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
b - Howard Hughes Medical Institute, *Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
c - Proteomics Platform, *Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA*
d - David H. Koch Institute for Integrative Cancer Research, *Massachusetts Institute of Technology, Barbara K. Ostrom Bioinformatics and Computing facility at the Swanson Biotechnology Center, Cambridge, MA 02139, USA*

Correspondence to : *A. Naba, Koch Institute for Integrative Cancer Research at MIT, Room 76-361, Cambridge, MA 02139, U.S.A. anaba@mit.edu. R.O. Hynes, Koch Institute for Integrative Cancer Research at MIT, Room 76-361D, Cambridge, MA 02139, U.S.A. rohynes@mit.edu*
http://dx.doi.org/10.1016/j.matbio.2015.06.003
Edited by R. Iozzo

Abstract

The extracellular matrix (ECM) is a fundamental component of multicellular organisms that provides mechanical and chemical cues that orchestrate cellular and tissue organization and functions. Degradation, hyperproduction or alteration of the composition of the ECM cause or accompany numerous pathologies. Thus, a better characterization of ECM composition, metabolism, and biology can lead to the identification of novel prognostic and diagnostic markers and therapeutic opportunities. The development over the last few years of high-throughput ("omics") approaches has considerably accelerated the pace of discovery in life sciences. In this review, we describe new bioinformatic tools and experimental strategies for ECM research, and illustrate how these tools and approaches can be exploited to provide novel insights in our understanding of ECM biology. We also introduce a web platform "the matrisome project" and the database MatrisomeDB that compiles *in silico* and *in vivo* data on the matrisome, defined as the ensemble of genes encoding ECM and ECM-associated proteins. Finally, we present a first draft of an ECM atlas built by compiling proteomics data on the ECM composition of 14 different tissues and tumor types.
© 2015 Elsevier B.V. All rights reserved.

Introduction

The extracellular matrix (ECM), a complex meshwork of proteins, is a fundamental component of multicellular organisms. The ECM is the three-dimensional architectural scaffold that defines tissue boundaries and biomechanical properties [1] and cell polarity. It also serves as an adhesive substrate for cell migration and, by binding morphogens and growth factors [2], can create concentration gradients for haptotactic migration [3,4] or pattern formation [5]. Extracellular matrix proteins provide biochemical cues interpreted by cell surface receptors, such as the integrins [6] and initiate signaling cascades controlling cell survival, cell proliferation, differentiation and stem cell state [7,8].

Advances in genome sequencing have allowed one to trace the evolution of the ECM and have revealed that, although some domains or modules characteristic of ECM proteins existed in unicellular organisms, the elaboration of ECM proteins and ECMs appeared largely in metazoa [9,10]. The multiplication and diversification of ECM proteins have accompanied major evolutionary innovations, in particular in the vertebrate phylum, including the appearance of a closed vascular system and of structures such as the neural crest, teeth, cartilage and bones [9,10]. During development, ECM plays vital roles in stem cell niches and in guiding migration and polarity of cells and axonal projections and in morphogenesis and coherence of tissues.

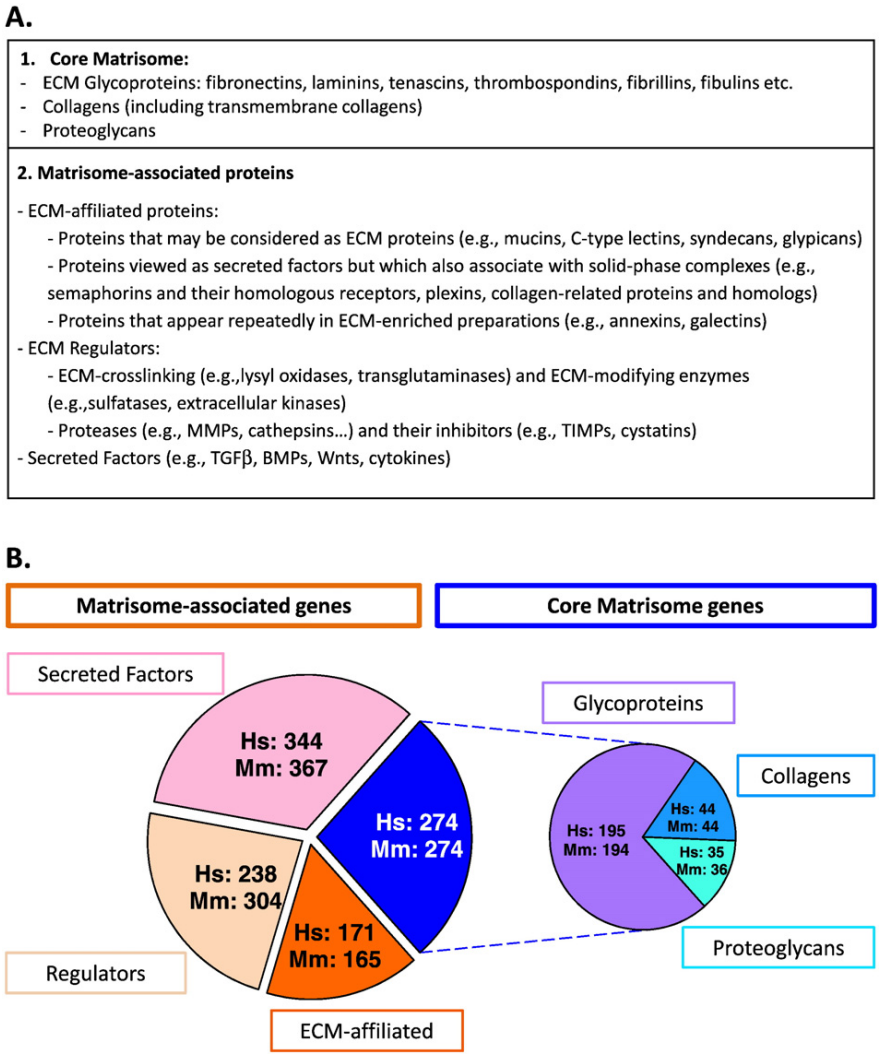

Fig. 1. Defining the matrisome. A. Definition of matrisome categories. The core matrisome comprises ECM glycoproteins, collagens and proteoglycans. Matrisome-associated proteins include ECM-affiliated proteins, ECM regulators and secreted factors [16,17]. B. Pie charts represent the number of human genes encoding core matrisome and matrisome-associated proteins. "Hs" indicates genes in the human genome and "Mm", genes in the murine genome.

Remodeling of ECM is essential during angiogenesis and branching morphogenesis of glands and in wound healing, as it is in cancer invasion [5].

Mutations in ECM genes are causal of musculoskeletal, cardio-vascular, renal, ocular and skin, diseases [11]. In addition, excessive deposition or, conversely, destruction of the ECM can also lead to pathologies such as fibrosis or osteoarthritis [12]. ECM deposition (or desmoplasia) is a hallmark of tumor progression and has been used by pathologists as a marker of tumors with poor prognosis even long before the composition and the complexity of the ECM were uncovered. Recent studies from us and others have revealed that the ECM plays a functional role in tumor progression and dissemination [13,14].

We previously proposed that characterizing the global composition of the extracellular matrix proteome, or "matrisome" of normal and diseased tissues would lead to important discoveries. This proposition raised two important questions: how do we define the extracellular matrix and how can one study the composition of a compartment largely made of insoluble proteins? Here, we review the bioinformatic tools and experimental approaches that have been developed over the last few years to characterize the composition of extracellular matrices. We also introduce a web interface "the matrisome project" that hosts a novel database on ECM and ECM-associated genes and proteins, MatrisomeDB that compiles *in silico* and experimental resources and we illustrate

their utility to analyze various "omics" datasets. Finally, we present here a first draft of an ECM atlas built by compiling proteomics data on the extracellular matrix of 14 different tissues and tumor types.

In silico definition of the matrisome

Bioinformatic approach to define the matrisome

In 1984, Martin and colleagues coined the term "matrisome" in the context of basement membranes to define "supramolecular complexes of matrix components which are the functional units of the forming extracellular matrix" [15]. In 2012, we proposed to extend the definition of this term to include not only *all* the genes encoding structural ECM components but also genes encoding proteins that can (or may) interact with or remodel the ECM [16–18] (Fig. 1A). To define the matrisome, we devised a bioinformatic approach to screen the human and mouse proteomes (or any other genomes of interest) using defining features of ECM proteins such as presence of signal peptide and of protein domains characteristic of ECM proteins [16,19]. We further curated manually the lists obtained computationally and, based on structural or functional features, distinguished core matrisome proteins from matrisome-associated proteins. The core matrisome comprises ECM glycoproteins, collagens and proteoglycans (Fig. 1A) (for reviews on each of these categories of proteins see [20–22]). ECM-associated proteins include ECM-affiliated proteins, ECM regulators and secreted factors that may interact with core ECM proteins (Fig. 1A) [16].

We note here that the criteria for assigning certain proteins to the matrisome categories can be debated. For example, the designation of the various secreted factors was deliberately inclusive since, although many of them are not known to associate with ECM, some clearly do and arguments can be adduced supporting the concept that many others do as well [2]. Equally, decisions such as where, or whether, to include some proteins containing short collagen triple helical domains or transmembrane segments are to some extent arbitrary. We view the current categorizations to be a working structure, subject to periodic revision in light of future discoveries.

Matrisome 2.0

The *in silico* definition of the matrisome relied on the interrogation of the protein database UniProt [23] using lists of domains from the InterPro [24] and SMART [25,26] databases. The protein-centric lists originally defined were then turned into gene-centric lists using NCBI Entrez Gene as reference gene database [27]. Databases are, by nature, very dynamic and constantly updated. The release of a major update by UniProt in July 2014 with improved identification of isoforms, new knowledge (such as the identification of the first extracellular kinases [28]) and discussions with colleagues have prompted us to update the original matrisome lists. This updated version of the matrisome (v2.0) comprises 1027 genes for the human genome and 1110 genes for the mouse genome (Fig. 1B). These numbers are slightly smaller than the ones we reported in 2012; this is mainly due to better annotations of pseudogenes (now removed from the matrisome lists) and removal of duplicate entries in the newer UniProt database. It is worth noting that the relatively large difference in the number of genes encoding ECM regulators in human and mouse is mostly due to gene duplication of proteases such as ADAMs (a disintegrin and metalloprotease domain proteins) and serpins (serine protease inhibitors) in the mouse genome. In addition to updating the gene-centric lists, we have also updated the list of UniProt accession numbers associated with each matrisome gene. These updated matrisome lists as well as updated domain lists are available for download from the web interface we recently developed http://matrisomeproject.mit.edu [29]. This website hosts a new database, MatrisomeDB, that provides, for all human and murine matrisome genes, live cross-referencing to gene (HUGO Gene Nomenclature Committee [30] and Mouse Genome Informatics [31]) and protein (UniProt, InterPro) databases, and information on gene orthology. In addition to providing links to *in silico* resources on matrisome genes and proteins, MatrisomeDB also incorporates experimental proteomics data generated in the Hynes lab. Database users can now easily determine whether or not a given ECM protein has been detected by proteomics in the tissues and tumors profiled (see below).

Matrisome lists as annotation tools for big data

Identifying ECM signatures in gene and protein datasets

We previously noted that the matrisome lists and categories we defined are significantly more comprehensive than Gene Ontology's "Cellular Component" categories for data mining and for posing questions relevant to ECM biology, since ECM proteins are currently scattered in multiple GO categories and comingled with non-ECM proteins [17]. In order to further facilitate the use of categorical matrisome lists for annotating genomic, transcriptomic or proteomic outputs and identifying ECM signatures

Table 1. Matrisome gene sets available in the Molecular Signature Database (MSigDB v5.0) and integrated in the collection C2: Canonical Pathways.
This table provides a description of the 10 matrisome gene sets and links to their pages in MSigDB where lists of genes can be downloaded.

Name	# genes	Description
NABA_MATRISOME	1027	Ensemble of genes encoding extracellular matrix and extracellular matrix-associated proteins
NABA_CORE_MATRISOME	274	Ensemble of genes encoding core extracellular matrix including ECM glycoproteins, collagens and proteoglycans
NABA_ECM_GLYCOPROTEINS	195	Genes encoding structural ECM glycoproteins
NABA_COLLAGENS	44	Genes encoding collagen proteins
NABA_PROTEOGLYCANS	35	Genes encoding proteoglycans
NABA_BASEMENT_MEMBRANES	40	Genes encoding structural components of basement membranes
NABA_MATRISOME_ASSOCIATED	753	Ensemble of genes encoding ECM-associated proteins including ECM-affiliated proteins, ECM regulators and secreted factors
NABA_ECM_AFFILIATED	171	Genes encoding proteins affiliated structurally or functionally to extracellular matrix proteins
NABA_ECM_REGULATORS	238	Genes encoding enzymes and their regulators involved in the remodeling of the extracellular matrix
NABA_SECRETED_FACTORS	344	Genes encoding secreted soluble factors

within datasets, we derived a collection of ten matrisome gene sets (Table 1) and implemented them into the Molecular Signatures Database maintained by the Broad Institute. In addition to allowing rapid annotation of large datasets and facilitating the identification of ECM signatures, these gene sets can now readily be used to conduct Gene Set Enrichment Analysis (GSEA) studies [32,33]. GSEA is a computational method that determines whether an *a priori* defined set of genes (for example, the matrisome) shows statistically significant, concordant differences between two biological states. Hussenet, Orend and collaborators used such an approach to identify a matrisome signature enriched in the transcriptome of angiogenic *vs* non-angiogenic oncogenic pancreatic islets, which they termed the "angiomatrix" signature [34]. Once identified, ECM signatures (of a given tissue or disease state) can be further analyzed with pathway or interaction analysis tools such as MatrixDB [35,36].

ECM signatures can also be used to interrogate other publicly available datasets. For example, we compared the ECM signatures of primary colorectal tumors and their liver metastases defined by proteomic analysis of small numbers of patient samples, with gene expression data from four clinical studies representing over 200 patients and demonstrated the association of a subset of the ECM proteins, defined by proteomics, with tumor progression [37].

Interrogating databases with matrisome lists

The two examples above illustrate the value of using limited experimental data on ECM genes or proteins to access and leverage the burgeoning pool of publicly available data from diverse modes of analysis. The development of high-throughput transcriptomics studies (microarray or RNAseq approaches) has been accompanied by the creation of databases to make the large body of data publicly available [38,39]. Recent papers have also reported the first drafts of the human proteome [40–42] which has led to the development of a website http://proteomicsdb.org. One can therefore use the matrisome list or sublists and ask which ECM genes or proteins are detected in which tissues, at which developmental stage(s), *etc*. As an example, we interrogated proteomicsDB with the list of 44 human collagen genes (Fig. 2). Although the most abundant and ubiquitous collagens (type I, III, IV, VI) were detected in most tissue and cell proteomes, less abundant or rarer collagens were seen only in a few tissues or in none so far. This may reflect reality but could also result from the fact that most studies included in proteomicsDB were not designed to look specifically for ECM proteins and may well have missed them. The very nature of ECM proteins, which are often very large, highly glycosylated, cross-linked, and difficult to solubilize, has made biochemical analyses of the composition of extracellular matrices challenging.

Experimental characterization of *in vivo* matrisomes by proteomics

Technical challenges and proposed solutions

ECM protein enrichment

The profiling of the protein composition of extracellular matrices by mass spectrometry is dependent on the availability of methods that selectively enrich for ECM proteins. Taking advantage of the insolubility of ECM proteins, we and others have reported the development of decellularization and ECM-enrichment methods that rely on the extraction of –

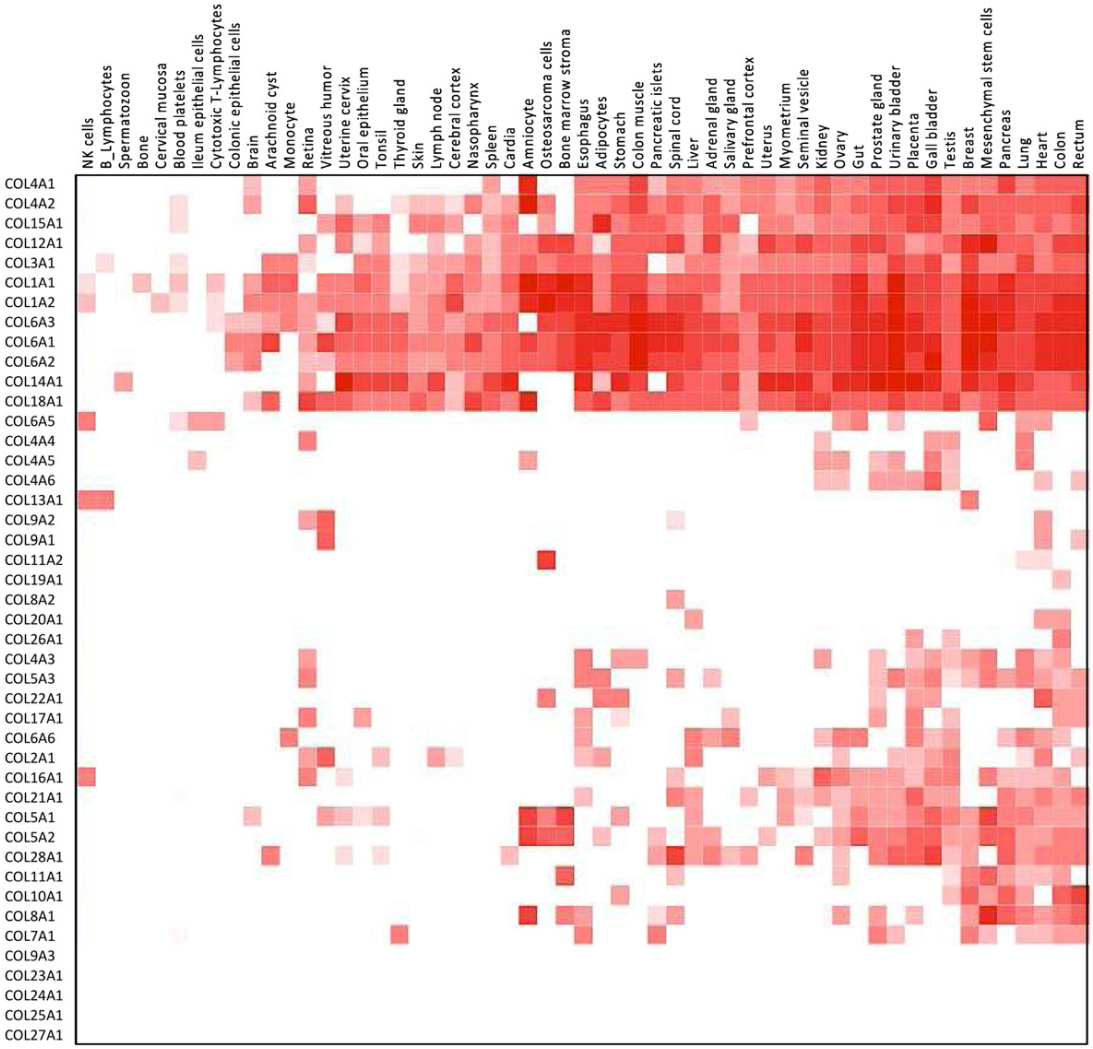

Fig. 2. Expression of collagen chains across 55 tissue and cellular proteomes. ProteomicsDB (April 2015 release) was interrogated with the 44 human collagen genes. The expression of 39 out of the 44 human collagen chains was reported in at least one of the 55 tissues or cell types for which global proteomics data were available. The five collagen chains not detected are: COL9A3, COL23A1, COL24A1, COL25A1, and COL27A1.

somewhat more soluble – intracellular proteins [16,43–46]. The insolubility of ECM proteins is, however, a challenge for subsequent proteomic analyses; indeed mass-spectrometry pipelines require proteins to be solubilized and then digested into peptides.

ECM protein solubilization

Complete ECM protein solubilization requires reagents in concentrations that are incompatible with mass spectrometry. In addition, a fraction of the extracellular matrix is resistant to solubilizing agents (such as SDS or urea) and would be lost if only solubilized material were carried forward. Therefore, we proposed that *crude* ECM-enriched samples be denatured, reduced and alkylated, deglycosylated, and digested by a combination of proteases (LysC and trypsin), all without intervening centrifugation, in order to minimize losses of otherwise insoluble materials. Using this approach, we could show that solubilization occurs as a concomitant of protease treatment [16,43].

Identification of MS spectra and peptides

ECM proteins, in particular collagens, are known to contain extensive sites of unique posttranslational modifications such as hydroxylated lysines or prolines [47,48]. It is thus important, when

Table 2. List of normal and diseased tissues, for which ECMs have been profiled by mass spectrometry.

Tissue_species	Species	References
Normal tissues		
Aorta	Human	Didangelos et al., 2010 [52]
Bone	Rat	Schrieweis et al., 2007 [50]
Cartilage; growth plate cartilage	Mouse	Belluoccio et al., 2006 [49]; Wilson et al., 2010 [54]
Articular and tracheal cartilages, meniscus, intervertebral disc, ribs	Human	Önnerfjord et al., 2012 [58]; Müller et al., 2014 [61]
Colon	Mouse	Naba et al., 2012 [16][a]
Colon	Human	Naba et al., 2014b [37][a]
Lung	Human	Booth et al., 2012 [57]
Lung	Rat	Hill et al., 2015 [63]
Liver	Human	Naba et al., 2014b [37][a]
Mammary gland	Rat	Hansen et al., 2009 [46]; O'Brien et al., 2012 [53]
Glomerular basement membrane	Human	Lennon et al., 2014 [60][a]
Retinal vascular basement membrane, Lens capsule, Inner limiting membrane	Human	Uechi et al., 2014 [62][a]
Retinal vascular basement membrane	Chick embryo	Balasubramani et al., 2010 [51]
Diseased tissues		
Poorly and highly metastatic melanoma xenografts	Human/mouse	Naba et al., 2012 [16][a]
Poorly and highly metastatic mammary carcinoma xenografts	Human/mouse	Naba et al., 2014a [13][a]
Metastatic primary colon carcinoma and derived liver metastases	Human	Naba et al., 2014b [37][a]
Fibrotic lung	Mouse	Decaris et al., 2014 [59]
Abdominal aortic aneurysm	Human	Didangelos et al., 2011 [55]
Cardiac ECM remodeling during ischemia/reperfusion	Pig	Barallobre-Barreiro et al., 2012 [56]

[a] Indicates studies for which raw mass spectrometry were made publicly available and used to build the draft of the ECM atlas (see Tables 3 & 4 and Supplementary Table 1).

conducting database searches for identification of spectra, to allow for such posttranslational modifications.

Using pipelines combining ECM-enrichment procedures and mass spectrometry, we and others have reported the characterization of the composition of ECMs of several normal and diseased tissues (Table 2) [16,13,37,46,49–63]. As expected, these studies revealed the presence of largely ubiquitously expressed ECM proteins but also identified tissue-specific proteins. For example the Önnerfjord laboratory showed that different cartilages have different ECM compositions [58]. Moreover, using an approach combining quantitative and targeted mass spectrometry, they showed that the composition of articular cartilage presents spatial variation and demonstrated that asporin, tenascin-C, thrombospondin-4 and perlecan were the most abundant in the superficial cartilage layer, whereas mimecan and thrombospondin-1 were most abundant in the intermediate layer and aggrecan, osteomodulin, chondroadherin were enriched in deeper layers of articular cartilage. The Mayr laboratory reported the characterization of the cardiac ECM and aortic basement membrane and further identified ECM signatures of ischemia/reperfusion and abdominal aortic aneurysm [52,55,56]. Using human melanoma and mammary carcinoma xenografts in mouse, and the ability of mass spectrometry to distinguish human (tumor-derived) protein sequences from their murine (stroma- or host-derived) counterparts, we have demonstrated that both the tumor cells and the stromal cells contribute to the production of the tumor ECM [13,16]. We have also shown that several ECM proteins differentially expressed between poorly and highly metastatic tumors (including Latent TGFβ Binding Protein 3, LTBP3, and the protein Sushi, Nidogen, and EGF-like Domains 1, SNED1) were causal of a more metastatic phenotype [13]. Furthermore, the comparison of the matrisomes of normal and paired tumor samples identified consistent differences in the ECMs of (i) primary tumors as compared with normal surrounding tissues, (ii) metastases and the normal tissue in which they develop, and (iii) primary tumors as compared with metastases derived from them [37]. Finally, using both xenografts and human-patient-derived samples, we could show that some ECM proteins identified could serve as potential biomarkers for tumor progression, metastasis and survival [13,37].

Building an ECM atlas

The large amount of data generated in "omics" studies has prompted researchers to share their data with the scientific community through dedicated repositories such as Gene Expression Omnibus or GEO for transcriptomics data and ProteomeXchange for proteomics data. In fact more and more

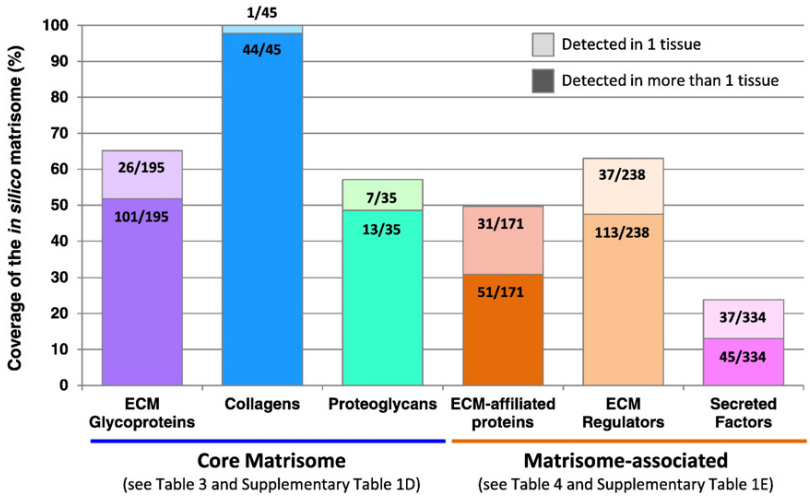

Fig. 3. Experimental coverage of the *in silico* of the matrisome. Bar chart represents, for each matrisome category, the percentage representation and number of ECM genes detected in one (lighter bars) or more than one (darker bars) tissue.

journals now strongly encourage or even require that the raw data accompanying publications be deposited in public databases.

We sought to build a first draft of an ECM atlas by integrating publicly available mass spectrometry data from studies designed specifically to characterize the global compositions of ECMs (Table 2). We gathered raw mass spectrometric data from our own group [13,16,37], data on the glomerular basement membrane from the Humphries lab [60] and data on three different ocular basement membranes (retinal vascular basement membrane, lens capsule, inner limiting membrane) from the Balasubramani group [62]. To provide consistent data analysis, all proteomics datasets were re-analyzed using the search pipeline we previously developed and peptides identified with a false discovery rate $<1.6\%$ were assembled into proteins (using a UniProt database) and annotated as being ECM-derived or not (Supplementary Table 1) [16]. As the data were acquired on different mass spectrometers, we cannot readily derive quantitative information regarding the relative abundance of each protein across tissues. Nonetheless, this compilation provides a first draft of an ECM atlas.

Compared with the "*in silico*" matrisome, this study shows that more than 60% of the predicted ECM glycoproteins were detected in one or more tissues (Fig. 3). Interestingly, several ECM proteins, in particular members of the insulin-like-growth-factor-binding protein family (IGFBP3, IGFBP4, IGFBP5), members of the CCN family (CTGF and CYR61), thrombospondin-2 (Thbs2), tenascin-N (Tnn) and VWA9 were only detected, so far, in ECM-enriched preparations from tumor samples (colorectal or mammary carcinomas, melanomas) but not in those from normal tissues (Table 3, Supplementary Table 1C, 1D). All the collagen chains were detected in one or more tissues, even though some are often thought to be cartilage-specific (type II and type IX collagens) (Fig. 3 and Table 3). Close to 60% of the predicted proteoglycans were detected. We hypothesize that the matrisome proteins not yet detected will be found in other tissues or at different developmental stages. For example, many ECM proteins are known to be exclusively expressed in teeth (such as ameloblastin, amelogenins, dentin sialophosphoprotein) [64], bones (such as bone gamma-carboxyglutamate (Gla) protein or osteocalcin, integrin-binding sialoprotein), the inner ear (such as the tectorins) or the nervous system [65,66].

ECM proteins exist in many different isoforms, and the presence or absence of certain spliced domains can change the function of a protein or be indicative of a diseased state. For example, fibronectin can include two fibronectin type III domains, each encoded by one exon (EIIIB and EIIIA), and the expression of the isoforms of fibronectin including one or both exons have been shown to be restricted to stages of development, remodeling tumor vasculature and sites of injuries [67]. Similar results are true for spliced variants of tenascins [68] and other ECM proteins. The protein database UniProt we used to generate this first draft of the ECM atlas includes comprehensive nomenclature and data on protein isoforms. We were thus able to provide isoform information for several of the core matrisome proteins constituting the ECM atlas (see Table 3 and Supplementary Table 1C and 1D).

The experimental coverage of the predicted matrisome-associated proteins is not as high: 50% of the ECM-affiliated proteins, over 60% of the ECM regulators and only 25% of the secreted factors have been detected in at least one tissue and a large

Table 3. **ECM atlas part I: Core matrisome proteins detected by mass spectrometry.**
List of ECM glycoproteins, collagens and proteoglycans detected with at least 2 peptides in any of the 14 tissues compiled in the ECM atlas: human glomerular basement membrane, human retinal vascular basement membrane, human inner limiting membrane (eye), human lens capsule, murine lung, murine colon, human colon, human liver, poorly and highly metastatic melanoma and mammary carcinoma xenografts. See Supplementary Table 1C and 1D for details.

ECM GLYCOPROTEINS	Glomerular basement membrane	Retinal blood vessel basement membrane	Inner limiting membrane	Lens capsule	Normal lung	Normal colon (mouse)	Normal colon (human)	Normal liver	Primary metastatic colon cancer	Liver metastasis from colon cancer	Poorly metastatic melanoma xenograft	Highly metastatic melanoma xenograft	Poorly metastatic mammary tumor xenograft	Highly metastatic mammary tumor xenograft
EMILIN1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
FN1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LAMA4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LAMA5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LAMB2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LAMC1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
NID1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
NID2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TGFBI	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TNC*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
VTN	■	■	■	■	■	■	■	■	■	■	■	■	■	■
VWA1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LAMA2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
FBN1*	■	■		■	■	■	■	■	■	■	■	■	■	
FBN2	■	■		■	■	■	■	■	■	■	■	■	■	■
FGA	■	■		■	■	■	■	■	■	■	■	■	■	■
FGG	■	■		■	■	■	■	■	■	■	■	■	■	■
LAMB1	■	■		■	■	■	■	■	■	■	■	■	■	■
TINAGL1*	■	■		■	■	■	■	■	■	■	■	■	■	■
TNXB*	■	■		■	■	■	■	■	■	■	■	■	■	■
AGRN*	■	■	■	■	■	■		■	■	■	■		■	■
LAMA3	■	■		■	■	■	■	■	■	■	■	■	■	■
FGB	■	■			■	■	■	■	■	■	■	■	■	■
Igfbp7	■	■			■	■	■	■	■	■	■	■	■	■
DPT	■			■	■	■	■	■	■	■	■	■	■	■
POSTN*	■			■	■	■	■	■	■	■	■	■	■	■
PXDN	■	■		■	■	■		■	■	■	■	■	■	■
FBLN5	■	■			■	■	■	■		■	■	■	■	■
Efemp1*	■				■	■	■	■	■	■	■	■	■	■
MFAP2		■		■	■	■	■	■	■	■	■		■	■
ECM1		■			■	■		■	■	■		■	■	■
LTBP2		■		■	■			■	■	■	■	■	■	■
ELN					■	■	■	■	■	■	■	■	■	■
EMILIN2					■	■	■	■	■	■	■	■	■	■
FBLN2					■	■	■	■	■	■	■	■	■	■
LTBP1*					■	■	■	■	■	■	■	■	■	■
Ltbp4*					■	■	■	■	■	■	■	■	■	■
LAMA1	■	■	■	■	■	■	■		■		■			
PAPLN*	■	■	■	■	■	■		■	■				■	
VWF		■		■	■	■	■	■	■		■			■
AEBP1		■		■		■	■	■	■	■	■		■	
HMCN1*		■			■	■	■	■	■	■	■			
THSD4*		■			■	■	■	■	■	■	■	■	■	■
SRPX*				■		■	■	■	■	■	■	■	■	■
MFGE8	■	■		■	■			■	■	■				■
Lamc2	■				■	■	■		■	■			■	■
THBS1	■				■				■	■	■	■	■	■
Mfap5					■	■	■	■	■	■			■	■
Vwa5a					■	■			■	■	■	■	■	■
MATN2	■	■		■		■	■		■		■			
HMCN2	■	■			■	■			■					■
MFAP4	■				■	■	■	■	■	■				
EMILIN3		■		■		■	■	■	■			■		
LAMB3					■	■	■		■	■			■	■
SBSPON	■	■	■	■	■					■				
LAMB4	■	■		■			■	■	■					
Emid1*	■	■			■	■	■		■					
CILP	■		■				■		■		■	■		
MMRN2	■			■	■			■	■		■			
FBLN1*	■				■		■	■	■	■				
Npnt*	■				■		■	■	■	■				
COLQ*	■					■	■	■	■		■			
ABI3BP*		■		■	■	■	■		■					
EFEMP2*					■			■	■	■	■			■
PCOLCE					■				■	■	■	■		■

ECM GLYCOPROTEINS	Glomerular basement membrane	Retinal blood vessel basement membrane	Inner limiting membrane	Lens capsule	Normal lung	Normal colon (mouse)	Normal colon (human)	Normal liver	Primary metastatic colon cancer	Liver metastasis from colon cancer	Poorly metastatic melanoma xenograft	Highly metastatic melanoma xenograft	Poorly metastatic mammary tumor xenograft	Highly metastatic mammary tumor xenograft
LTBP3*						■			■	■	■	■		■
IGFBP3									■	■			■	■
MXRA5	■	■					■		■	■				
TINAG	■				■	■			■	■				
LAMC3	■				■		■		■	■				
SPON1		■	■	■	■				■					
SRPX2						■	■		■	■			■	
MGP	■	■		■					■	■				
GLDN	■				■				■					■
FNDC1	■					■			■	■				
SNED1	■										■	■		■
LGI4		■		■			■		■					
MMRN1					■	■		■	■					
FGL2							■	■	■	■				
THBS2									■	■	■		■	
FRAS1	■	■			■									
NTN4	■			■	■									
IGFALS	■							■		■				
Dmbt1		■				■			■					
FBN3		■							■	■				
OTOG			■			■			■					
Thbs3					■	■					■			
Sparc					■				■					■
Wisp2					■			■		■				
SVEP1								■	■	■				
VWA9									■		■			■
NTN3		■	■											
FBLN7		■		■										
Matn4					■	■								
Cilp2					■		■							
Bmper					■			■						
IGFBP5									■	■				
IGFBP4										■				■
Tnn											■	■		
CTGF													■	■
CYR61													■	■
FNDC8	■													
MFAP3	■													
RELN	■													
VWCE	■													
LGI1		■												
MEPE		■												
MFAP1		■												
PXDNL		■												
SPARCL1		■												
Igfbp6						■								
Kcp						■								
Matn1						■								
Ndnf						■								
Ntn1						■								
Pcolce2						■								
Thbs4							■							
CRELD1									■					
DSPP									■					
ECM2									■					
VWA5B1									■					
COMP										■				
SPON2										■				
SPP1*										■				
TNFAIP6										■				
VWDE													■	
AW551984														■

* Indicates proteins for which more than one isoform was detected.

Searchable tables are provided in Supplementary Table 1.

COLLAGENS	Glomerular basement membrane	Retinal blood vessel basement membrane	Inner limiting membrane	Lens capsule	Normal lung	Normal colon (mouse)	Normal colon (human)	Normal liver	Primary metastatic colon cancer	Liver metastasis from colon cancer	Poorly metastatic melanoma xenograft	Highly metastatic melanoma xenograft	Poorly metastatic mammary tumor xenograft	Highly metastatic mammary tumor xenograft
COL11A2*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL14A1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL18A1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL1A1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL1A2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL22A1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL2A1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL3A1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL4A1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL4A2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL4A3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL4A5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL5A1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL5A2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL6A2*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL6A3*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL4A4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL4A6*	■	■	■	■	■	■	■	■	■	■	■	■	■	
COL9A3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COL12A1	■	■		■	■	■	■	■	■	■	■	■	■	■
COL5A3	■	■		■	■	■	■	■	■	■	■	■	■	■
COL6A1	■	■		■	■	■	■	■	■	■	■	■	■	■
COL7A1	■	■		■	■	■	■	■	■	■	■	■	■	■
COL8A1	■	■		■	■	■	■	■	■	■	■	■		■
COL10A1	■	■			■	■	■	■	■	■	■	■	■	■
COL11A1	■	■			■	■	■	■	■	■	■	■	■	■
COL19A1	■	■			■	■	■	■	■	■	■	■	■	■
COL16A1	■		■		■	■	■	■	■	■	■	■	■	■
COL15A1	■			■	■	■	■	■	■	■	■	■	■	■
COL9A1	■	■	■	■	■	■	■	■	■	■	■			■
COL27A1*	■	■		■	■	■	■	■	■	■	■	■		■
COL24A1	■	■		■	■	■		■	■	■	■	■		■
COL9A2		■	■	■	■	■	■	■			■	■	■	■
COL23A1*	■	■			■	■	■	■	■	■	■	■		
COL13A1*	■				■	■	■	■	■		■	■	■	■
COL8A2*	■				■	■	■	■	■	■	■	■	■	■
Col6a5	■			■	■	■		■		■	■	■	■	
Col6a6	■			■	■	■		■	■		■	■	■	
COL25A1*	■				■	■	■	■	■	■				
COL28A1					■	■	■	■	■	■	■		■	■
COL21A1*	■	■				■	■	■	■	■		■		■
COL17A1	■					■	■		■	■	■			■
COL26A1	■						■	■						
COL20A1*	■						■		■					
Col6a4						■								

PROTEOGLYCANS	Glomerular basement membrane	Retinal blood vessel basement membrane	Inner limiting membrane	Lens capsule	Normal lung	Normal colon (mouse)	Normal colon (human)	Normal liver	Primary metastatic colon cancer	Liver metastasis from colon cancer	Poorly metastatic melanoma xenograft	Highly metastatic melanoma xenograft	Poorly metastatic mammary tumor xenograft	Highly metastatic mammary tumor xenograft
BGN	■	■	■	■	■	■	■	■	■	■	■	■	■	■
HSPG2*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LUM	■	■		■	■	■	■	■	■	■	■	■	■	■
PRELP	■	■		■	■	■	■	■	■	■	■	■	■	■
ASPN	■			■	■	■	■	■	■	■	■	■	■	■
DCN	■			■	■	■	■	■	■	■	■	■	■	■
VCAN	■		■		■	■	■	■	■	■	■	■	■	
OGN				■	■	■	■	■	■	■	■		■	
PRG4	■			■		■		■	■	■	■	■		
PRG2		■			■	■		■	■	■	■			
PRG3								■	■	■	■			
FMOD								■		■	■	■		
CHADL		■	■	■										
OPTC		■		■										
IMPG1		■		■	■									
PODN				■					■	■				
HAPLN1						■						■		
Acan					■									
Chad					■									
Spock3						■								

Table 4. **ECM atlas part II: Matrisome-associated proteins detected by mass spectrometry.**
List of ECM-affiliated proteins, ECM regulators and secreted factors detected with at least 2 peptides in any of the 14 tissues compiled in the ECM atlas: human glomerular basement membrane, human retinal vascular basement membrane, human inner limiting membrane (eye), human lens capsule, murine lung, murine colon, human colon, human liver, poorly and highly metastatic melanoma and mammary carcinoma xenografts. See Supplementary Table 1C and 1E for details.

ECM REGULATORS	Glomerular basement membrane	Retinal blood vessel basement membrane	Inner limiting membrane	Lens capsule	Normal lung	Normal colon (mouse)	Normal colon (human)	Normal liver	Primary metastatic colon cancer	Liver metastasis from colon cancer	Poorly metastatic melanoma xenograft	Highly metastatic melanoma xenograft	Poorly metastatic mammary tumor xenograft	Highly metastatic mammary tumor xenograft
SERPINH1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TGM2*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
AMBP*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TIMP3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
F13A1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PZP	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSD	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLG	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPING1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ITIH2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Loxl1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ITIH5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSB	■	■	■	■	■	■	■	■	■	■	■	■	■	■
KNG1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Itih1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINE2*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
HTRA1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
F2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSG	■	■	■	■	■	■	■	■	■	■	■	■	■	■
A2M	■	■	■	■	■	■	■	■	■	■	■	■	■	■
F9	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMTSL1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ITIH4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMTSL4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LOXL2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLOD1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
HRG	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINC1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Serpinf2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLOD3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CSTB	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINB1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ITIH3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LOXL3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Serpina3k	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINB6	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PRSS1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ELANE	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP9	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINA1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINA3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAM9	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINF1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Serpind1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINB12	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LOXL4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAM10	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Lox	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LEPRE1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP14	■	■	■	■	■	■	■	■	■	■	■	■	■	■
P4HA1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLOD2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP12	■	■	■	■	■	■	■	■	■	■	■	■	■	■
AGT	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TGM3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TGM1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
BMP1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TGM5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Mug2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Serpina1a	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Serpina1b	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSS	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TIMP1*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINB9	■	■	■	■	■	■	■	■	■	■	■	■	■	■
F10	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PAPPA	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMTSL5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP23B*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CST3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CD109	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINB3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C17orf58	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Serpina3m	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Serpina1d	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ST14	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMDEC1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
F12	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSA	■	■	■	■	■	■	■	■	■	■	■	■	■	■
F13B	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LEPREL2*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSZ	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSF	■	■	■	■	■	■	■	■	■	■	■	■	■	■

ECM REGULATORS	Glomerular basement membrane	Retinal blood vessel basement membrane	Inner limiting membrane	Lens capsule	Normal lung	Normal colon (mouse)	Normal colon (human)	Normal liver	Primary metastatic colon cancer	Liver metastasis from colon cancer	Poorly metastatic melanoma xenograft	Highly metastatic melanoma xenograft	Poorly metastatic mammary tumor xenograft	Highly metastatic mammary tumor xenograft
SERPINB8	■	■	■	■	■	■	■	■	■	■	■	■	■	■
NGLY1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLAU	■	■	■	■	■	■	■	■	■	■	■	■	■	■
EGLN1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
P4HA2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CST6	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LPA	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAM8*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP7	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINA5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CSTA	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINB4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP28	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SULF1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAM19	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Mmp19	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Serpina3n	■	■	■	■	■	■	■	■	■	■	■	■	■	■
FAM20B	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Ky	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SLPI	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAM15	■	■	■	■	■	■	■	■	■	■	■	■	■	■
HTRA3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP11	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMTS12	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSH	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINA4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
HABP2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINA10	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TLL1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAM12	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMTS13	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ITIH6	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMTSL2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSK	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSV	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PRSS12	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMTS16	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMTS5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
1810010H24Rik	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Adamts17	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Adamts9	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Hyal2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Adam7	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Leprel1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Mep1a	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Serpinb1a	■	■	■	■	■	■	■	■	■	■	■	■	■	■
FAM20C	■	■	■	■	■	■	■	■	■	■	■	■	■	■
HPSE2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAMTS10	■	■	■	■	■	■	■	■	■	■	■	■	■	■
HPSE	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP10	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MMP3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINB7	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINB10	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINB5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TLL2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSL	■	■	■	■	■	■	■	■	■	■	■	■	■	■
F7	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Adam23	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SERPINE1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Try10	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLAT	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ADAM17	■	■	■	■	■	■	■	■	■	■	■	■	■	■
AI182371	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CTSC	■	■	■	■	■	■	■	■	■	■	■	■	■	■
P4HTM	■	■	■	■	■	■	■	■	■	■	■	■	■	■

* Indicates proteins for which more than one isoform was detected.

Searchable tables are provided in Supplementary Table 1.

ECM-AFFILIATED PROTEINS	Glomerular basement membrane	Retinal blood vessel basement membrane	Inner limiting membrane	Lens capsule	Normal lung	Normal colon (mouse)	Normal colon (human)	Normal liver	Primary metastatic colon cancer	Liver metastasis from colon cancer	Poorly metastatic melanoma xenograft	Highly metastatic melanoma xenograft	Poorly metastatic mammary tumor xenograft	Highly metastatic mammary tumor xenograft
ANXA1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ANXA2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ANXA6*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ANXA11	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LGALS1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LGALS3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LMAN1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LGALS9	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ANXA5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ANXA7	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CSPG4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1QC	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Sftpd	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ANXA4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LGALS8	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1QB	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLXNB2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SEMA3B	■	■	■	■	■	■	■	■	■	■	■	■	■	■
FCN3*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC16	■	■	■	■	■	■	■	■	■	■	■	■	■	■
HPX	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ANXA3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LGALS4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1QA	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLXND1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Colec12	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Sema3c	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CLEC18A	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Frem2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1QL3	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1QL2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1QTNF5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GPC4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Pixdc2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC5B	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CLEC11A	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SEMA6D	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CSPG5	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SEMA7A	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SEMA4C	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CLEC18B	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LGALS7	■	■	■	■	■	■	■	■	■	■	■	■	■	■
COLEC11*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC13	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLXNC1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC4*	■	■	■	■	■	■	■	■	■	■	■	■	■	■
REG4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ANXA13	■	■	■	■	■	■	■	■	■	■	■	■	■	■
FCN1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GREM1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1QL4	■	■	■	■	■	■	■	■	■	■	■	■	■	■
LGALS2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC20	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC6	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SEMA3F	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SEMA4D	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Anxa9	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Clec14a	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Frem1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Itln1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Sftpa1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Sftpb	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Sftpc	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1qtnf9	■	■	■	■	■	■	■	■	■	■	■	■	■	■
ITLN2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MBL2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GPC1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
GPC6	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC12	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLXDC1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLXNA1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
PLXNB1	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1QTNF2	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CLC	■	■	■	■	■	■	■	■	■	■	■	■	■	■
CLEC12A	■	■	■	■	■	■	■	■	■	■	■	■	■	■
MUC19	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SEMA3A	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SEMA4A	■	■	■	■	■	■	■	■	■	■	■	■	■	■
SEMA4B	■	■	■	■	■	■	■	■	■	■	■	■	■	■
C1qtnf7	■	■	■	■	■	■	■	■	■	■	■	■	■	■

| SECRETED FACTORS | Glomerular basement membrane | Retinal blood vessel basement membrane | Inner limiting membrane | Lens capsule |

stoichiometry and their low abundance may compromise their identification by mass spectrometry. Some proteins apparently specific to a given tissue type may simply be present but in too low abundance in other tissues to be detected. This may be overcome by implementing "targeted" mass spectrometry approaches, which allow focus on the measurement of a subset of proteins suspected or known to be present in a given sample [69,70]. Matrisome-associated proteins are also likely to be more soluble than core matrisome proteins and could be lost during decellularization. One approach to retrieve information on more soluble proteins would be to profile not only the composition of the insoluble ECM fraction generated by tissue decellularization but also the composition of the intermediate fractions generated during the decellularization process [43,63] and the soluble components of tissue interstitial fluid. Of course, these intermediate fractions will also be enriched for intracellular components, thus robust data annotations using matrisome lists could assist with delineating which proteins are soluble matrisome proteins and which are contaminants. Finally, when we initially defined the list of matrisome-associated proteins, we wanted to be inclusive and have included entire families of proteins some of which may not be found in close association with core ECM proteins. This may have resulted in an over-prediction and, based on future experiments, we may revise our definition and distinguish more precisely those proteins belonging to the matrisome from those belonging to the secretome (the human protein atlas [71] predicts that over 3000 human genes have a secreted product).

Future directions in ECM research

With the tools and methods now in place, we can postulate and certainly hope that the completion of a human ECM atlas will be achievable within the next few years. We would like to incorporate additional datasets (including the ones listed in Table 2) in the atlas if/when they become publicly accessible and we would like to encourage the deposition of all future research data in the public domain.

Time- and spatially-resolved matrisomes

The analysis of the composition of the ECM of tissues has already revealed novel or unsuspected components of the ECM characteristics of a diseased state or causal of a diseased state. Efforts should now focus on obtaining temporally- and spatially-resolved matrisomes. For those interested in the role of the ECM in diseases, we propose that quantitative proteomics will become a method of choice to profile the dynamic changes in the composition of the extracellular matrix that occur during disease progression or during the course of treatment [72]. We hypothesize that this type of study will allow the identification of novel prognostic or diagnostic ECM markers that could assist clinical decisions (see below).

Development of high-throughput approaches to map post-translational modifications of ECM proteins

ECM proteins undergo extensive post-translational modifications including hydroxylation, phosphorylation, sulfation, glycosylation, and crosslinking, that can have significant physiological and/or pathological implications [73]. We propose that beyond the profiling of the ECM proteome, the community should aim for broad implementation of novel "omics" approaches such as glycomics [74–76] to profile the post-translational modifications of ECM proteins.

ECM proteins also undergo proteolytic cleavage and release fragments as part of their physiological (or pathological) turnover. Increased ECM degradation is a hallmark of pathologies such as osteoarthritis, fibrosis, and cancers [12]. ECM protein fragments displaying biological activities are termed matricryptins or matrikines [77]. These fragments are characterized by novel amino- and/or carboxy-terminal extremities and can thus be identified by mass spectrometry. The emergence of degradomics or terminomics approaches [78] offers interesting opportunities for ECM research. In addition, it has been proposed that the identification of cleaved fragments can also serve as proxy for identifying proteases (e.g. MMPs, cathepsins) active in a given tissue [79].

Translational applications

ECM proteins as biomarkers

Alterations of the extracellular matrix are responsible for, or accompany, the development of pathologies, such as skeletal and articular diseases, cardiovascular diseases, skin diseases, fibrosis and cancers. We have previously shown that comparison of the matrisomes of normal and cancerous tissues can lead to the identification of novel candidate biomarkers [37]. We thus propose that pursuing the effort of characterizing the composition of extracellular matrices could lead to the identification of novel biomarkers for other diseases in addition to cancers. It is worth noting that ECM proteins are particularly favorable candidate biomarkers for immunohistochemically based assays since they are readily accessible, abundant, and laid down in characteristic patterns. Once identified, disease-specific ECM proteins, or protein isoforms could also serve as anchors for imaging molecules (e.g. fluorescent molecules, radiotracers) or therapeutics (e.g. drugs, cytokines, radioisotopes) that could be coupled, for example, to anti-disease-

specific ECM protein antibodies, on the model of systems developed in the Neri lab [80,81]. In addition, proteolytic fragments of ECM proteins can be released in body fluids and could be used as readouts for disease progression or treatment efficiency [82–84].

ECM proteins as therapeutic targets

There are successful precedents for inhibiting integrins to treat thrombosis, auto-immune and inflammatory diseases, *etc.* [85,86]. However, there are also many unsuccessful examples, including the use of matrix metalloproteinase inhibitors, in clinical trials for cancer patients [87]. In a 2009 review, Järveläinen and colleagues noted that ECM proteins are often ignored in drug discovery efforts [88]. We postulate that this will change and that high-throughput approaches will permit the identification of novel disease-specific ECM proteins and protein isoforms and that the characterization of the molecular mechanisms downstream of these proteins will offer novel therapeutic opportunities. Therapeutic strategies could include (i) targeting ECM protein synthesis or post-translational modifications, (ii) targeting ECM remodeling (degradation or crosslinking) and (iii) targeting ECM/ECM receptor interactions [88] as well as the antibody-mediated targeting discussed above.

ECM and regenerative medicine

The ECM provides biophysical and biochemical signals that orchestrate organ formation and function and thus should play a central part in tissue engineering strategies [89]. In addition, the ECM is a fundamental component of stem cell niches [90,91]. In fact, many stem cell markers are ECM receptors (e.g. CD49a-f are integrins α1-6, CD29 is the integrin β1 and Lgr5 is an R-spondin receptor) and laminins have been demonstrated to affect pluripotency and differentiation of stem cells [90]. Tissue engineers have exploited, to varying degrees, aspects of ECM biology, in particular its biomechanical properties (not discussed here), to design novel scaffolds to support tissue regeneration [92]. Allogeneic or xenogeneic ECMs are now being used routinely to aid the reconstruction of a variety of tissues (e.g. urinary bladder, skin) [93]. Whole organ engineering, if successful, should solve the problem of shortage of organs available to patients awaiting transplant. One approach to engineer a whole functional organ (heart, lung) consists in using decellularized organ scaffolds comprising ECM and some (generally unknown) associated materials and repopulating the scaffold with stem cells. However, so far this approach has not yet succeeded in regenerating fully functional organs. This may be in part due to the fact that decellularization results, as shown by mass spectrometry, in the loss of some ECM components and associated growth factors [63,94,95]. At the opposite extreme, minimalistic approaches are parsimoniously incorporating features of ECM proteins such as RGD (integrin-binding) peptides or mimics in artificial scaffolds. However, proteomic studies have revealed that the ECMs of tissues are made of 150+ proteins and although reconstructing this complexity may be difficult (and perhaps unnecessary [96]), we propose that the results of proteomics studies aimed at characterizing *in vivo* ECMs should be exploited to guide the design of the next generation of bio-inspired scaffolds to support organ regeneration.

Conclusions

The extracellular matrix has long been considered as a structural component of tissue organization. Recognition of the roles of specific ECM receptors and the binding and presentation of secreted growth factors introduced new concepts of how ECM proteins affect cell behavior and the developing understanding of mechanochemical transduction of ECM-derived signals have all combined to implicate ECM in a wide range of biologically and medically important areas [97]. With the recent development of experimental techniques for thorough characterization of ECM composition and of bioinformatic means to exploit that information, we are at an exciting point in ECM research where we can deploy the power of broad-scale genomic, proteomic and other “omic” approaches to provide new insights into development, disease, therapy and regenerative medicine in addition to further fundamental understanding of the enduring fascinating mysteries of ECM biology.

Acknowledgments

The authors would like to thank Drs. Jill Mesirov and Arthur Liberzon from the Broad Institute's Computational Biology and Bioinformatics Organization for implementing the matrisome gene sets into the Molecular Signature Database.

The authors would also like to thank the laboratories of Drs. Martin Humphries (University of Manchester) and Manimalha Balasubramani (University of Pittsburgh) for making their raw mass spectrometry data publicly available.

This work was supported by grants from the National Cancer Institute — Tumor Microenvironment Network (U54 CA126515/CA163109), the Broad Institute of MIT and Harvard, the Howard Hughes Medical Institute, of which ROH is an Investigator and in part by Support Grant to the Koch Institute (P30-CA14051 from the National Cancer Institute). AN was supported by postdoctoral

fellowships from the Howard Hughes Medical Institute and the Ludwig Center for Cancer Research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.matbio.2015.06.003.

Received 9 May 2015;
Received in revised form 9 June 2015;
Accepted 10 June 2015
Available online 8 July 2015

Keywords:
Extracellular matrix;
Matrisome;
MatrisomeDB;
ECM atlas;
Bioinformatics;
Proteomics

References

[1] J.D. Humphrey, E.R. Dufresne, M.A. Schwartz, Mechanotransduction and extracellular matrix homeostasis, Nat. Rev. Mol. Cell Biol. 15 (2014) 802–812, http://dx.doi.org/10.1038/nrm3896.
[2] R.O. Hynes, The extracellular matrix: not just pretty fibrils, Science 326 (2009) 1216–1219, http://dx.doi.org/10.1126/science.1176009.
[3] A. Huttenlocher, A.R. Horwitz, Integrins in cell migration, Cold Spring Harb. Perspect. Biol. 3 (2011) a005074, http://dx.doi.org/10.1101/cshperspect.a005074.
[4] G. Charras, E. Sahai, Physical influences of the extracellular environment on cell migration, Nat. Rev. Mol. Cell Biol. (2014)http://dx.doi.org/10.1038/nrm3897.
[5] D.E. Ingber, Extracellular matrix: a solid-state regulator of cell form, function, and tissue development, Comprehensive Physiology 2011, pp. 541–556.
[6] I.D. Campbell, M.J. Humphries, Integrin structure, activation, and interactions, Cold Spring Harb. Perspect. Biol. (2011)http://dx.doi.org/10.1101/cshperspect.a004994.
[7] T. Rozario, D.W. DeSimone, The extracellular matrix in development and morphogenesis: a dynamic view, Dev. Biol. 341 (2010) 126–140, http://dx.doi.org/10.1016/j.ydbio.2009.10.026.
[8] S.A. Wickström, K. Radovanac, R. Fässler, Genetic analyses of integrin signaling, Cold Spring Harb. Perspect. Biol. (2011)http://dx.doi.org/10.1101/cshperspect.a005116.
[9] R.O. Hynes, The evolution of metazoan extracellular matrix, J. Cell Biol. 196 (2012) 671–679, http://dx.doi.org/10.1083/jcb.201109041.
[10] F.W. Keeley, R.P. Mecham (Eds.), Evolution of Extracellular Matrix, Springer, Berlin, Heidelberg, 2013.
[11] J.F. Bateman, R.P. Boot-Handford, S.R. Lamandé, Genetic diseases of connective tissues: cellular and extracellular effects of ECM mutations, Nat. Rev. Genet. 10 (2009) 173–183, http://dx.doi.org/10.1038/nrg2520.
[12] C. Bonnans, J. Chou, Z. Werb, Remodelling the extracellular matrix in development and disease, Nat. Rev. Mol. Cell Biol. 15 (2014) 786–801, http://dx.doi.org/10.1038/nrm3904.
[13] A. Naba, K.R. Clauser, J.M. Lamar, et al., Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters, eLife (2014)http://dx.doi.org/10.7554/eLife.01308.
[14] M.W. Pickup, J.K. Mouw, V.M. Weaver, The extracellular matrix modulates the hallmarks of cancer, EMBO Rep. 15 (2014) 1243–1253, http://dx.doi.org/10.15252/embr.201439246.
[15] G.R. Martin, H.K. Kleinman, V.P. Terranova, et al., The regulation of basement membrane formation and cell-matrix interactions by defined supramolecular complexes, Ciba Found. Symp. 108 (1984) 197–212.
[16] A. Naba, K.R. Clauser, S. Hoersch, et al., The matrisome: *in silico* definition and *in vivo* characterization by proteomics of normal and tumor extracellular matrices, Mol. Cell. Proteomics 11 (2012)http://dx.doi.org/10.1074/mcp.M111.014647 (M111.014647).
[17] A. Naba, S. Hoersch, R.O. Hynes, Towards definition of an ECM parts list: an advance on GO categories, Matrix Biol. 31 (2012) 371–372, http://dx.doi.org/10.1016/j.matbio.2012.11.008.
[18] R.O. Hynes, A. Naba, Overview of the matrisome—an inventory of extracellular matrix constituents and functions, Cold Spring Harb. Perspect. Biol. (2012)http://dx.doi.org/10.1101/cshperspect.a004903.
[19] E. Hohenester, J. Engel, Domain structure and organisation in extracellular matrix proteins, Matrix Biol. 21 (2002) 115–128, http://dx.doi.org/10.1016/S0945-053X(01)00191-3.
[20] R.O. Hynes, K.M. Yamada (Eds.), Extracellular Matrix Biology., Cold Spring Harbor Perspectives in Biology, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2012.
[21] R.V. Iozzo, L. Schaefer, Proteoglycan form and function: a comprehensive nomenclature of proteoglycans, Matrix Biol. 42 (2015) 11–55, http://dx.doi.org/10.1016/j.matbio.2015.02.003.
[22] R. Mecham, The Extracellular Matrix: an Overview, Springer Science & Business Media, 2011.
[23] Consortium TU, UniProt: a hub for protein information, Nucleic Acids Res. 43 (2015) D204–D212, http://dx.doi.org/10.1093/nar/gku989.
[24] A. Mitchell, H.-Y. Chang, L. Daugherty, et al., The InterPro protein families database: the classification resource after 15 years, Nucleic Acids Res. 43 (2015) D213–D221, http://dx.doi.org/10.1093/nar/gku1243.
[25] J. Schultz, F. Milpetz, P. Bork, C.P. Ponting, SMART, a simple modular architecture research tool: identification of signaling domains, Proc. Natl. Acad. Sci. 95 (1998) 5857–5864.
[26] I. Letunic, T. Doerks, P. Bork, SMART: recent updates, new developments and status in 2015, Nucleic Acids Res. 43 (2015) D257–D260, http://dx.doi.org/10.1093/nar/gku949.
[27] D. Maglott, J. Ostell, K.D. Pruitt, T. Tatusova, Entrez Gene: gene-centered information at NCBI, Nucleic Acids Res. 39 (2011) D52–D57, http://dx.doi.org/10.1093/nar/gkq1237.
[28] V.S. Tagliabracci, J.L. Engel, J. Wen, et al., Secreted kinase phosphorylates extracellular proteins that regulate biomineralization, Science 336 (2012) 1150–1153, http://dx.doi.org/10.1126/science.1217817.
[29] A. Naba, H. Ding, C.A. Whittaker, R.O. Hynes, http://matrisomeproject.mit.edu.

[30] K.A. Gray, L.C. Daugherty, S.M. Gordon, et al., Genenames.org: the HGNC resources in 2013, Nucleic Acids Res. 41 (2013) D545–D552, http://dx.doi.org/10.1093/nar/gks1066.
[31] Jackson Laboratories Mouse Genome Informatics website., http://www.informatics.jax.org/.
[32] A. Liberzon, A description of the Molecular Signatures Database (MSigDB) web site, Methods Mol. Biol. Clifton, NJ 1150 (2014) 153–160, http://dx.doi.org/10.1007/978-1-4939-0512-6_9.
[33] A. Subramanian, P. Tamayo, V.K. Mootha, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. 102 (2005) 15545–15550, http://dx.doi.org/10.1073/pnas.0506580102.
[34] B. Langlois, F. Saupe, T. Rupp, et al., AngioMatrix, a signature of the tumor angiogenic switch-specific matrisome, correlates with poor prognosis for glioma and colorectal cancer patients, Oncotarget 5 (2014) 10529–10545.
[35] E. Chautard, L. Ballut, N. Thierry-Mieg, S. Ricard-Blum, MatrixDB, a database focused on extracellular protein–protein and protein–carbohydrate interactions, Bioinformatics 25 (2009) 690–691, http://dx.doi.org/10.1093/bioinformatics/btp025.
[36] G. Launay, R. Salza, D. Multedo, et al., MatrixDB, the extracellular matrix interaction database: updated content, a new navigator and expanded functionalities, Nucleic Acids Res. 43 (2015) D321–D327, http://dx.doi.org/10.1093/nar/gku1091.
[37] A. Naba, K.R. Clauser, C.A. Whittaker, et al., Extracellular matrix signatures of human primary metastatic colon cancers and their metastases to liver, BMC Cancer 14 (2014) 518, http://dx.doi.org/10.1186/1471-2407-14-518.
[38] M. Kapushesky, T. Adamusiak, T. Burdett, et al., Gene Expression Atlas update—a value-added database of microarray and sequencing-based functional genomics experiments, Nucleic Acids Res. 40 (2012) D1077–D1081, http://dx.doi.org/10.1093/nar/gkr913.
[39] R. Petryszak, T. Burdett, B. Fiorelli, et al., Expression Atlas update—a database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments, Nucleic Acids Res. 42 (2014) D926–D932, http://dx.doi.org/10.1093/nar/gkt1270.
[40] M.-S. Kim, S.M. Pinto, D. Getnet, et al., A draft map of the human proteome, Nature 509 (2014) 575–581, http://dx.doi.org/10.1038/nature13302.
[41] M. Wilhelm, J. Schlegl, H. Hahne, et al., Mass-spectrometry-based draft of the human proteome, Nature 509 (2014) 582–587, http://dx.doi.org/10.1038/nature13319.
[42] M. Uhlén, L. Fagerberg, B.M. Hallström, et al., Tissue-based map of the human proteome, Science 347 (2015) 1260419, http://dx.doi.org/10.1126/science.1260419.
[43] A. Naba, K.R. Clauser, R.O. Hynes, Enrichment of extracellular matrix proteins from tissues and digestion into peptides for mass spectrometry analysis, J. Vis. Exp. (2015)http://dx.doi.org/10.3791/53057.
[44] J. Barallobre-Barreiro, A. Didangelos, X. Yin, et al., A sequential extraction methodology for cardiac extracellular matrix prior to proteomics analysis, in: F. Vivanco (Ed.), Heart Proteomics, Humana Press, Totowa, NJ 2013, pp. 215–223.
[45] R. Wilson, D. Belluoccio, J.F. Bateman, Proteomic analysis of cartilage proteins, Methods 45 (2008) 22–31, http://dx.doi.org/10.1016/j.ymeth.2008.01.008.
[46] K.C. Hansen, L. Kiemele, O. Maller, et al., An in-solution ultrasonication-assisted digestion method for improved extracellular matrix proteome coverage, Mol. Cell. Proteomics 8 (2009) 1648–1657, http://dx.doi.org/10.1074/mcp.M900039-MCP200.
[47] J. Myllyharju, Intracellular post-translational modifications of collagens, in: J. Brinckmann, H. Notbohm, P.K. Müller (Eds.), Collagen, Springer Berlin Heidelberg 2005, pp. 115–147.
[48] M. Yamauchi, M. Sricholpech, Lysine post-translational modifications of collagen, Essays Biochem. 52 (2012) 113–133, http://dx.doi.org/10.1042/bse0520113.
[49] D. Belluoccio, R. Wilson, D.J. Thornton, et al., Proteomic analysis of mouse growth plate cartilage, Proteomics 6 (2006) 6549–6553, http://dx.doi.org/10.1002/pmic.200600191.
[50] M.A. Schreiweis, J.P. Butler, N.H. Kulkarni, et al., A proteomic analysis of adult rat bone reveals the presence of cartilage/chondrocyte markers, J. Cell. Biochem. 101 (2007) 466–476, http://dx.doi.org/10.1002/jcb.21196.
[51] M. Balasubramani, E.M. Schreiber, J. Candiello, et al., Molecular interactions in the retinal basement membrane system: a proteomic approach, Matrix Biol. 29 (2010) 471–483, http://dx.doi.org/10.1016/j.matbio.2010.04.002.
[52] A. Didangelos, X. Yin, K. Mandal, et al., Proteomics characterization of extracellular space components in the human aorta, Mol. Cell. Proteomics 9 (2010) 2048–2062, http://dx.doi.org/10.1074/mcp.M110.001693.
[53] J. O'Brien, J. Fornetti, P. Schedin, Isolation of mammary-specific extracellular matrix to assess acute cell-ECM interactions in 3D culture, J. Mammary Gland Biol. Neoplasia 15 (2010) 353–364, http://dx.doi.org/10.1007/s10911-010-9185-x.
[54] R. Wilson, A.F. Diseberg, L. Gordon, et al., Comprehensive profiling of cartilage extracellular matrix formation and maturation using sequential extraction and label-free quantitative proteomics, Mol. Cell. Proteomics 9 (2010) 1296–1313, http://dx.doi.org/10.1074/mcp.M000014-MCP201.
[55] A. Didangelos, X. Yin, K. Mandal, et al., Extracellular matrix composition and remodeling in human abdominal aortic aneurysms: a proteomics approach, Mol. Cell. Proteomics 10 (2011)http://dx.doi.org/10.1074/mcp.M111.008128 (M111.008128).
[56] J. Barallobre-Barreiro, A. Didangelos, F.A. Schoendube, et al., Proteomics analysis of cardiac extracellular matrix remodeling in a porcine model of ischemia/reperfusion injury, Circulation 125 (2012) 789–802, http://dx.doi.org/10.1161/CIRCULATIONAHA.111.056952.
[57] A.J. Booth, R. Hadley, A.M. Cornett, et al., Acellular normal and fibrotic human lung matrices as a culture system for *in vitro* investigation, Am. J. Respir. Crit. Care Med. 186 (2012) 866–876, http://dx.doi.org/10.1164/rccm.201204-0754OC.
[58] P. Önnerfjord, A. Khabut, F.P. Reinholt, et al., Quantitative proteomic analysis of eight cartilaginous tissues reveals characteristic differences as well as similarities between subgroups, J. Biol. Chem. 287 (2012) 18913–18924, http://dx.doi.org/10.1074/jbc.M111.298968.
[59] M.L. Decaris, M. Gatmaitan, S. FlorCruz, et al., Proteomic analysis of altered extracellular matrix turnover in bleomycin-induced pulmonary fibrosis, Mol. Cell. Proteomics 13 (2014) 1741–1752, http://dx.doi.org/10.1074/mcp.M113.037267.
[60] R. Lennon, A. Byron, J.D. Humphries, et al., Global analysis reveals the complexity of the human glomerular extracellular matrix, J. Am. Soc. Nephrol. 25 (2014) 939–951, http://dx.doi.org/10.1681/ASN.2013030233.
[61] C. Müller, A. Khabut, J. Dudhia, et al., Quantitative proteomics at different depths in human articular cartilage reveals unique

patterns of protein distribution, Matrix Biol. 40 (2014) 34–45, http://dx.doi.org/10.1016/j.matbio.2014.08.013.
[62] G. Uechi, Z. Sun, E.M. Schreiber, et al., Proteomic view of basement membranes from human retinal blood vessels, inner limiting membranes, and lens capsules, J. Proteome Res. (2014)http://dx.doi.org/10.1021/pr5002065.
[63] R.C. Hill, E.A. Calle, M. Dzieciatkowska, et al., Quantification of extracellular matrix proteins from a rat lung scaffold to provide a molecular readout for tissue engineering, Mol. Cell. Proteomics MCP 14 (2015) 961–973, http://dx.doi.org/10.1074/mcp.M114.045260.
[64] J.C.-C. Hu, Y. Yamakoshi, F. Yamakoshi, et al., Proteomics and genetics of dental enamel, Cells Tissues Organs 181 (2005) 219–231, http://dx.doi.org/10.1159/000091383.
[65] D.R. Zimmermann, M.T. Dours-Zimmermann, Extracellular matrix of the central nervous system: from neglect to challenge, Histochem. Cell Biol. 130 (2008) 635–653, http://dx.doi.org/10.1007/s00418-008-0485-9.
[66] C.S. Barros, S.J. Franco, U. Müller, Extracellular matrix: functions in the nervous system, Cold Spring Harb. Perspect. Biol. (2011)http://dx.doi.org/10.1101/cshperspect.a005108.
[67] J.E. Schwarzbauer, D.W. DeSimone, Fibronectins, their fibrillogenesis, and *in vivo* functions, Cold Spring Harb. Perspect. Biol. (2011)http://dx.doi.org/10.1101/cshperspect.a005041.
[68] R. Chiquet-Ehrismann, R.P. Tucker, Tenascins and the importance of adhesion modulation, Cold Spring Harb. Perspect. Biol. (2011)http://dx.doi.org/10.1101/cshperspect.a004960.
[69] E.S. Boja, H. Rodriguez, Mass spectrometry-based targeted quantitative proteomics: achieving sensitive and reproducible detection of proteins, Proteomics 12 (2012) 1093–1110, http://dx.doi.org/10.1002/pmic.201100387.
[70] S.A. Carr, S.E. Abbatiello, B.L. Ackermann, et al., Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based asay development using a fit-for-purpose approach, Mol. Cell. Proteomics 13 (2014) 907–917, http://dx.doi.org/10.1074/mcp.M113.036095.
[71] F. Pontén, K. Jirström, M. Uhlen, The Human Protein Atlas—a tool for pathology, J. Pathol. 216 (2008) 387–393, http://dx.doi.org/10.1002/path.2440.
[72] M.A. Gillette, S.A. Carr, Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry, Nat. Methods 10 (2013) 28–34, http://dx.doi.org/10.1038/nmeth.2309.
[73] G. Yalak, B.R. Olsen, Proteomic database mining opens up avenues utilizing extracellular protein phosphorylation for novel therapeutic applications, J. Transl. Med. 13 (2015) 125, http://dx.doi.org/10.1186/s12967-015-0482-4.
[74] R. Sasisekharan, R. Raman, V. Prabhakar, Glycomics approach to structure–function relationships of glycosaminoglycans, Annu. Rev. Biomed. Eng. 8 (2006) 181–231, http://dx.doi.org/10.1146/annurev.bioeng.8.061505.095745.
[75] M. Lech, I. Capila, G.V. Kaundinya, Mass spectrometric methods for the analysis of heparin and heparan sulfate, Methods Mol. Biol. Clifton NJ 1229 (2015) 119–128, http://dx.doi.org/10.1007/978-1-4939-1714-3_12.
[76] Y. Chiu, R. Huang, R. Orlando, J.S. Sharp, GAG-ID: Heparan Sulfate and Heparin Glycosaminoglycan High-Throughput Identification Software, Mol. Cell. Proteomics MCP (2015)http://dx.doi.org/10.1074/mcp.M114.045856.
[77] S. Ricard-Blum, R. Salza, Matricryptins and matrikines: biologically active fragments of the extracellular matrix, Exp. Dermatol. 23 (2014) 457–463, http://dx.doi.org/10.1111/exd.12435.
[78] L.D. Rogers, C.M. Overall, Proteolytic post-translational modification of proteins: proteomic tools and methodology, Mol. Cell. Proteomics 12 (2013) 3532–3542, http://dx.doi.org/10.1074/mcp.M113.031310.
[79] P. Schlage, U. Auf dem Keller, Proteomic approaches to uncover MMP function, Matrix Biol. 44–46 (2015) 232–238.
[80] M. Santimaria, G. Moscatelli, G.L. Viale, et al., Immunoscintigraphic detection of the ED-B domain of fibronectin, a marker of angiogenesis, in patients with cancer, Clin. Cancer Res. 9 (2003) 571–579.
[81] N. Pasche, D. Neri, Immunocytokines: a novel class of potent armed antibodies, Drug Discov. Today 17 (2012) 583–590, http://dx.doi.org/10.1016/j.drudis.2012.01.007.
[82] L. Yao, W. Lao, Y. Zhang, et al., Identification of EFEMP2 as a serum biomarker for the early detection of colorectal cancer with lectin affinity capture assisted secretome analysis of cultured fresh tissues, J. Proteome Res. 11 (2012) 3281–3294, http://dx.doi.org/10.1021/pr300020p.
[83] R. Schierwagen, D.J. Leeming, S. Klein, et al., Serum markers of the extracellular matrix remodeling reflect antifibrotic therapy in bile-duct ligated rats, Front. Physiol. 4 (2013) 195, http://dx.doi.org/10.3389/fphys.2013.00195.
[84] K. Bunatova, M. Pesta, V. Kulda, et al., Plasma TIMP1 level is a prognostic factor in patients with liver metastases, Anticancer Res. 32 (2012) 4601–4606.
[85] S.L. Goodman, M. Picard, Integrins as therapeutic targets, Trends Pharmacol. Sci. 33 (2012) 405–412, http://dx.doi.org/10.1016/j.tips.2012.04.002.
[86] B. Estevez, B. Shen, X. Du, Targeting integrin and integrin signaling in treating thrombosis, Arterioscler. Thromb. Vasc. Biol. 35 (2015) 24–29, http://dx.doi.org/10.1161/ATVBAHA.114.303411.
[87] R.E. Vandenbroucke, C. Libert, Is there new hope for therapeutic matrix metalloproteinase inhibition? Nat. Rev. Drug Discov. 13 (2014) 904–927, http://dx.doi.org/10.1038/nrd4390.
[88] H. Järveläinen, A. Sainio, M. Koulu, et al., Extracellular matrix molecules: potential targets in pharmacotherapy, Pharmacol. Rev. 61 (2009) 198–223, http://dx.doi.org/10.1124/pr.109.001289.
[89] T. Dvir, B.P. Timko, D.S. Kohane, R. Langer, Nanotechnological strategies for engineering complex tissues, Nat. Nanotechnol. 6 (2011) 13–22, http://dx.doi.org/10.1038/nnano.2010.246.
[90] F.M. Watt, W.T.S. Huck, Role of the extracellular matrix in regulating stem cell fate, Nat. Rev. Mol. Cell Biol. 14 (2013) 467–473, http://dx.doi.org/10.1038/nrm3620.
[91] F. Gattazzo, A. Urciuolo, P. Bonaldo, Extracellular matrix: a dynamic microenvironment for stem cell niche, Biochim. Biophys. Acta 1840 (2014) 2506–2519, http://dx.doi.org/10.1016/j.bbagen.2014.01.010.
[92] S.F. Badylak, D.J. Weiss, A. Caplan, P. Macchiarini, Engineered whole organs and complex tissues, Lancet 379 (2012) 943–952, http://dx.doi.org/10.1016/S0140-6736(12)60073-7.
[93] D.M. Faulk, S.A. Johnson, L. Zhang, S.F. Badylak, Role of the extracellular matrix in whole organ engineering, J. Cell. Physiol. 229 (2014) 984–989, http://dx.doi.org/10.1002/jcp.24532.
[94] R.W. Bonvillain, S. Danchuk, D.E. Sullivan, et al., A nonhuman primate model of lung regeneration: detergent-mediated decellularization and initial In vitro recellularization with mesenchymal stem cells, Tissue Eng. Part A 18 (2012) 2437–2452, http://dx.doi.org/10.1089/ten.tea.2011.0594.

[95] T.H. Petersen, E.A. Calle, L. Zhao, et al., Tissue-engineered lungs for *in vivo* implantation, Science 329 (2010) 538–541, http://dx.doi.org/10.1126/science.1189345.
[96] K.A. Kyburz, K.S. Anseth, Synthetic mimics of the extracellular matrix: how simple is complex enough? Ann. Biomed. Eng. 1–12 (2015)http://dx.doi.org/10.1007/s10439-015-1297-4.
[97] R.O. Hynes, Stretching the boundaries of extracellular matrix research, Nat. Rev. Mol. Cell Biol. 15 (2014) 761–763, http://dx.doi.org/10.1038/nrm3908.