Title: Multidimensional extracellular-matrix analysis reveals a silent pre-clinical signature in inflammatory disease

Authors: Elee Shimshoni[1], Ran Afik[1], Anjana Shenoy[2], Miri Adler[3], Luca Puricelli[4], Veronica Ghini[5], Odelia Mouhadeb[6,8], Nathan Gluck[6,8], Sigal Fishman[6,8], Lael Werner[7,8], Dror S. Shouval[7,8], Chen Varol[6,8], Alessandro Podestà[4], Paola Turano[5], Tamar Geiger[2], Paolo Milani[4], Claudio Luchinat[2], Uri Alon[3] and Irit Sagi[1*]

Affiliations:
[1]Department of Biological Regulation, Weizmann Institute of Science, Rehovot, Israel
[2]Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv-Yafo, Israel
[3]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel
[4]CIMAINA and Department of Physics, Università degli Studi di Milano, Milano, Italy
[5]CERM and Department of Chemistry, University of Florence, Italy
[6]Research Center for Digestive Tract and Liver Diseases, Tel Aviv Sourasky Medical Center, Tel Aviv-Yafo, Israel
[7]Pediatric Gastroenterology Unit, Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Tel Hashomer, Israel
[8]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv-Yafo, Israel

*Correspondence to: irit.sagi@weizmann.ac.il, Department of Biological Regulation, Weizmann Institute of Science, Rehovot, IL-76100, Israel. Tel : +972-8-9342130

Abstract: Extracellular-matrix remodeling is an emerging hallmark of complex tissue pathologies, generating biomechanical and molecular signals to the cells at various stages of disease progression. As a fundamental component of virtually all tissues, the extracellular matrix has been widely studied. But as a complete, defined and dynamic entity in its own right, it remains poorly understood. Here, we took a systematic, holistic approach, to investigate the multidimensional extracellular-matrix landscape, spanning structure, stiffness, and protein and metabolite compositions, both in healthy animals, and as a dynamic test case, during intestinal inflammation. By integrating computational and advanced imaging analyses, we correlate proteomic composition to structural damage while revealing the potential of the extracellular matrix as a metabolite reservoir. Unexpectedly, this comprehensive analysis uncovered a previously undescribed, silent pathological tissue state occurring before the onset of inflammatory symptoms in the colon, characterized by discrete morphological and molecular changes in the extracellular matrix, along with protein and biomarker variation. This early pathology is exhibited as a multidimensional extracellular-matrix footprint by two different types of colitis models and represents a tissue state that readily leads to full-blown disease. We offer this powerful diagnostic modality as a novel method, which we term “Multi-ECM”, and suggest that application of Multi-ECM to other tissue-damaging diseases will reveal additional clinically silent states that foretell their onset.

Introduction

The extracellular matrix (ECM) comprises a large portion of any mammalian tissue. Cells are born in the ECM, die in the ECM, and communicate with each other through the ECM[1,2]. This intricate mesh of fibrillary proteins, proteoglycans, glycoproteins and other glycosaminoglycans (GAGs) includes both the interstitial matrix and the basement membrane that underlies epithelial tissues, and is constantly remodeled by tightly regulated remodeling enzymes, such as matrix metalloproteinases (MMPs), lysyl oxidases and heparanases[3,4]. Importantly, the involvement of aberrant ECM remodeling, in terms of changes in the expression or activity of specific ECM components or enzymes, has been demonstrated in inflammation, by promoting immune cell infiltration and signaling[5–11].

However, the study of tissue pathology in autoimmune and inflammatory diseases is almost exclusively immunological and cell-centric, with already limited study of the ECM relegated to background or contextual consideration. The regard for the ECM as a supporting player is a consequence of the rather limited available techniques specifically designed for ECM analysis, along with the difficulty of applying techniques that are more common for the study of intracellular proteins or soluble signaling molecules to the insoluble ECM components, which are secreted by a variety of cellular sources to the same extracellular space and regulated on multiple levels. This creates a large knowledge gap between what has been attained using more classical tools to study ECM on the one hand, and on the other hand, the high resolution afforded by state-of-the-art immunological techniques.

Previous studies, including from our lab, have challenged this bystander role for the ECM, demonstrating its dynamics and participation in diseases such as cancer[12–15] and infections[16], as well as the role ECM remodeling plays in inflammation[10,17,18]. Based on

our hypothesis that ECM may be vitally active in both health and disease processes, we have undertaken a systematic upgrade to the analysis of the ECM and pathology, to propel the study of the ECM into a more central place in understanding of tissue dynamics, which we propose is warranted. Our previous work and work by others underscored the need for an integrative, multidimensional, ECM-centric approach that is not based on a single component or enzyme to the analysis of complex diseases.

Tissue damage manifesting in the ECM is a hallmark of inflammatory bowel diseases (IBD) progression and severity[5,19–23]. Nevertheless, these diseases have not been comprehensively analyzed from the viewpoint of the ECM, as a complete and defined entity, though the roles of some individual ECM enzymes and components have been studied[7,24–36]. Taking the largely overlooked importance of the ECM in certain pathologies as our point of departure, we have developed an integrative, multidimensional, ECM-centric methodology for analysis of complex tissue states, which we have named "Multi-ECM". This highly sensitive diagnostic paradigm for interrogation of ECM dynamics encompasses four dimension-metrics: 1) proteomic composition 2) structure 3) stiffness and 4) ECM-derived metabolites. Application of these measurements is novel for the study of colitis, and through this rigorous technical advance, we reveal the ability of the ECM to act as a metabolite reservoir, which to our knowledge, we are the first to report. Via an integrative analysis of these four dimensions, we also meticulously define a clinically silent pathological tissue state occurring before the onset of inflammatory symptoms[37] which we call the "pre-symptomatic ECM scaffold". This state is demonstrated both in mice at a pre-clinical time-point of an acute dextran-sodium sulfate (DSS)-induced colitis model and in healthy IL-10 mice, which are prone to develop spontaneous colitis, and is significantly different from that of healthy WT mice, as well as from that of acutely or

chronically ill mice. Importantly, as a consequence of analysis of this state using these four dimension-metrics, we conclude that no one of these approaches on its own would be sufficient to conclusively pinpoint and characterize this disease-prone state. This showcases the importance of our integrative ECM-centric approach both as a basis for novel clinical diagnostic applications and as a major advance in the laboratory study of the ECM as a dynamic entity.

Finally, building upon the potential translational relevance of this discovery in mouse, we investigated some molecular parameters of mouse "pre-symptomatic ECM scaffold" in human disease tissue, and find that features identified in mouse models using Multi-ECM are also observed in clinical isolates of inflammatory bowel disease in human. These results showcase the importance of our integrative ECM-centric approach and analysis, and highlight the ECM as a central player in defining healthy and diseased tissue state. We propose that the ECM is, itself, a potent driver of tissue state.

Results

Clinical characterization and timeline of murine colitis models

The study of IBD has benefited greatly from the use of murine models, representing different types of intestinal inflammation as a result of immune dysregulation and/or injury. To begin, we use a common murine model for IBD – the DSS-induced colitis model[38], which, with the appropriate calibration, results in a well-defined time-frame of acute colitis development in wild-type (WT) mice. We induced acute colitis in wild type (WT) C57BL/6 mice via addition of 1.25% DSS into the drinking water for seven days. Disease progression was monitored by body weight measurement, endoscopic evaluation and histological analysis of colonic tissue, in accordance with commonly used clinically diagnostic methods[39] at different time points – day 0 (healthy WT), day 4 and day 10. Degree of inflammation was scored during endoscopic

examination according to five parameters: mucosal transparency, mucosal wall granularity, stool consistency, blood vessel deformation and fibrin appearance in the colon lumen. Each parameter was given a score between 0-3, with 0 indicating healthy and 3 indicating maximally pathological, as previously reported[40]. An overall score of 0-4 was defined as "healthy"; 5-7 was defined as "mildly inflamed"; 8-11 as "inflamed"; and 12-15 as "severely inflamed". Colon inflammation under these conditions peaks between days 7-10, as detected by endoscopy (**Fig. 1A**). Indeed, at day 10 all mice present with the two highest categories of colon inflammation, with just over 75% reaching "severely inflamed" (**Fig. 1C**), with a corresponding high histological score (see Materials and Methods for details) of 9 out of 14 on average (**Fig. 1E**). Importantly, while clinical symptoms begin to appear after day 5, on day 4 most animals (over 90%) are defined as "healthy" (**Fig. 1C**), since they do not lose weight (**Fig. 1D**), appear healthy by endoscopic evaluation, and do not display inflammation that is consistently detectable via histopathological analysis (**Fig. 1B** and **1E**). Therefore, we regard day 4 as the "pre-symptomatic" state.

Since our goal was to use our ECM analysis to define the landscape of intestinal inflammation, including a range of positions along the health-disease axis, we also examined C57BL/6 IL-10$^{-/-}$ mice ("IL-10") mice that spontaneously develop chronic colon inflammation[41]. In these mice, inflammation is not chemically induced, but rather is thought to result from immune dysregulation, and phenotypically, they more closely resemble the clinical manifestation of chronic ulcerative colitis (UC) and Crohn's disease (CD) patients and mimic a monogenic form of IBD[42]. In order to synchronize and accelerate the onset of chronic colitis, IL-10 mice were exposed to 200ppm piroxicam in their food for 14 days [43]. The result of this exposure was the development of IBD-like symptoms, which can be detected and scored using

colonoscopy similarly to the well-established DSS-induced model (**Fig. 1A**). Histopathology in this model shows prominent immune cell infiltration into the mucosa, but with little ulceration or colonic crypt abolition, and therefore has a different scoring system than DSS (see Materials and Methods for details) (**Fig. 1B**). Importantly, healthy naïve IL-10 mice do not display histopathological parameters of inflammation, while ill mice have a high histopathological score of 5.5 out of 8 on average (**Fig. 1H**) The penetrance of disease in this model is different than in the DSS-induced inflammation, as just over 70% of the mice develop colitis at varying degrees, with a ~20% mortality rate (**Fig. 1F**). IL-10 mice lose weight only in early phases of colitis development, and the weight is regained as chronic colitis is established. For comparing acute (day 10 of DSS) to chronic inflammation, we selected a stable chronic inflammation time point for further investigation (day 28), when mice appear to be at, or surpass, their initial weight before treatment (**Fig. 1G**). For comparison to the pre-symptomatic day 4 time point in the DSS model, we chose to use naïve eight to twelve-week-old IL-10 mice that were deemed "healthy" by endoscopic evaluation. Taken together, pre-symptomatic states and clear disease states and disease initiating parameters were defined in two animal models: one for transient acute colitis (DSS) and the other for persistent chronic colitis (IL10).

ECM structural changes precede inflammatory symptoms

Having established pre-disease and disease state conditions, our first dimension of Multi-ECM was characterization of the inflammatory states in ECM structure and morphology by direct visualization. To do this, we first utilized second-harmonic generation (SHG) microscopy, which allows visualization of fibrillar collagen, without labeling, in native tissues. More detailed and high-resolution inspection of collagen and other ECM structural components, such as the basement membrane, was achieved using scanning electron microscopy (SEM) on decellularized

155 colon tissue from the same mice. Despite the downfalls of sample preparation, i.e., loss of certain components due to decellularization and artifacts that may be introduced by dehydration, EM provide us with high resolution images of the spatial organization of the ECM like no other imaging modality. SHG microscopy revealed that in the colon of the healthy WT mouse, the collagen (**Fig. 2A** in red) uniformly circumscribes the well-ordered colonic crypts, which appear
160 to be of even size and distribution, as also reported previously[5]. High resolution SEM reveals that these healthy crypts are uniformly covered by a dense mesh layer, which is most likely the basement membrane (**Fig. 2B**). Strikingly, both imaging modalities reveal that acute inflammation (day 10 of DSS model) has a profound effect on ECM structure, with destruction of crypt architecture, ECM degradation in certain regions and build up in others (**Fig. S1A**),
165 along with altered and heterogeneous crypt size (**Fig. 2A** and **2C**). Similarly, SHG imaging of chronic inflammation in IL-10 mice reveals deterioration of collagen structure supporting the crypt wall, making crypt boundaries less clear. Inter-crypt space, as well as their diameters, increases and collagen ultrastructure is uneven compared to the healthy IL-10 mice (**Fig. 2A** and **2C** and **Fig. S1A**). Further inspection of the ECM structure using SEM reinforced this finding:
170 naïve IL-10 crypt wall ECM appears fibrous, whereas in chronically ill IL-10 mice, the crypt wall ECM appears disintegrated and perforated, with thinner fibrils (**Fig. 2B** and **Fig. S1B**). Remarkably, at the pre-symptomatic state of the DSS model (day 4), remodeling events of the colonic ECM can also be observed, specifically, as thickening or deterioration of some crypt walls and changes in crypt diameter, where was observed with SHG microscopy (**Fig. 2A** and
175 **Fig. S1A**). Even more prominent ECM remodeling can be seen at this stage, using high-resolution SEM, as exposed fibrillar collagen on the inner crypt walls (**Fig. 2B**), due to

destruction of the top layer coating this interstitial ECM (**Fig. S1B**), which is mostly likely the basement membrane, according to its mesh-like morphology [44].

Most surprising was the comparison of the healthy IL-10 mice to their WT counterpart and to DSS day 4. We observe a strong resemblance in ECM structure between the colon of IL-10 mice and that of the pre-symptomatic, day 4 DSS model, with the characteristic exposure of fibrils on the crypt wall, as pointed out by arrows in **Fig. 2B**. To test whether this observation is phenomenologically robust, we quantified the frequency of fibril exposure in the two pre-symptomatic states (day 4 of DSS and healthy IL-10), as a percentage of crypts that have exposed walls, by comparing to the frequency in healthy WT mice. We found that this structural change occurs in both models at a statistically significant rate (**Fig. 2D**) compared to healthy WT. Thus, IL-10 mice, which are prone to developing chronic inflammation, develop a "leaky" ECM layer underlying the epithelium as a chronic, pre-symptomatic state, compared to the dense mesh covering the WT murine colon (**Fig. 2B** and **Fig. S1B**).

Taken together, the structural similarity of the ECM between the two murine pre-symptomatic states, leads us to hypothesize that exposed fibrillar collagen due to the deterioration of the layer that coats it, is a feature of what we term the "pre-symptomatic ECM scaffold". We propose this as a general phenomenon, or global biomarker, in colitis, revealing a silent inflammatory state, i.e., a tissue state that appears macroscopically normal but which already possesses structural lesions detectable using high-resolution microscopy.

Pre-symptomatic structural changes coincide with gelatinase activity

The gelatinases MMP-2 and MMP-9 primarily degrade basement membranes[45,46]. Therefore, having observed that the most prominent structural ECM alteration occurring at pre-symptomatic states involves perforation of the mesh-like structure underlying the epithelium and

200 covering collagen fibrils, which is most likely the basement membrane, we set to examine the presence of gelatinolytic activity in these tissues. First, we tested mRNA expression of both gelatinases, as well as other MMPs and the tissue inhibitor of MMPs (TIMP)-1. We found that both MMP-2 and -9 were significantly unregulated both at day 4 of the DSS model and in healthy IL-10 mice compared to healthy WT, while other enzymes shows a weaker trend and
205 TIMP-1, which inhibits MMP-9 was unchanged (**Fig. 2E**). Since the final regulatory step of MMPs occurs outside of the cell, with their activation, we used the activity-based assay, in situ zymography[47] to find the localization of active gelatinases in the tissue. Our analysis reveals elevated gelatinolytic activity in the mucosa of the pre-symptomatic states, compared to healthy WT (**Fig. 2F-G**).

210 This finding shows that while MMP-9, especially, is an established marker of IBD disease activity[29,48], it may also be a marker of pre-clinical states. Together with the structural analysis in the previous section, we show that elevated gelatinolytic activity is part of the structural alterations that we observe at pre-symptomatic states.

Colitis is associated with a gradual loss of colonic ECM rigidity

215 Having defined a difference in the ECM in this pre-symptomatic state, we applied a second dimension of Multi-ECM, which interrogates the changes in mechanical properties of the ECM. To do this, we used mesoscale atomic-force microscopy (AFM) to analyze the biomechanical properties of decellularized ECM scaffolds derived from colons of WT mice from both experimental models, using both pre-symptomatic (day 4 of DSS and healthy IL-10) and
220 pathological (day 10 of DSS and chronically ill IL-10) states, in addition to healthy WT.

We used micron-sized probes, instead of the classical sharp AFM tips (characterized by a radius of curvature of a few nm), to apply pressure to a 500mmX500mm ECM scaffold. This

modification permits acquisition of the scaffolds' mesoscopically averaged, statistically robust Young's Modulus (YM) values [49,50], which describes the collective contributions of nanoscale molecular components organized in micrometer-sized (70µmX70µm) domains in the ECM (see Materials and Methods for more details). Distributions of the measured YM values of colonic ECM samples are broad (**Fig. S2**), demonstrating the diversity of elastic properties within the same sample, and suggesting variability within the complex structure and composition of the ECM of each individual. Each mode in the distribution (**Fig. S2** and **Table S2**) represents the elastic properties of a region of the ECM with lateral dimensions (and thickness) of several microns, which we may define a structural and functional domain of the ECM. We term these The median values and average median values of the YM of these mechanical domains are reported in **Fig. 3A-B**.

In all cases, we observe a trend in which the ECMs of inflamed colons are generally softer than those of healthy colons, as indicated by the median YM (**Fig. 3A-B** and **Table S1**). Importantly, the structurally modified ECM in the pre-symptomatic state (day 4) of the DSS model has mechanical domains with an intermediate rigidity, with values between those of healthy and acutely inflamed states of DSS. The ECM of the chronically ill IL-10 mice is much softer than that of the healthy IL-10 mice, with a lower median YM of their mechanical domains (**Fig. 3B** and **Table S1**). Chronically ill IL-10 mice possess a YM value that is approximately two-fold lower than that of the acutely inflamed colon (**Fig. 3A-B** and **Table S1**).

Unexpectedly, we find that the rigidity of the ECM is similar in healthy IL-10 and healthy WT mice, both possessing relatively stiff colonic ECM, with median YM values of about 11.0kPa. This similarity is in contrast to clear structural differences between the two mice. These

245 results suggest a second silent pre-symptomatic state in colon pathology: one in which there is a detectably altered structure without reduction in ECM rigidity.

Altogether, our two-pronged analysis highlights a gradual softening of the ECM over the course of acute colitis development, until a peaking of acute symptoms which are accompanied by massive ECM tissue damage, considered a hallmark of disease progression. We find that loss
250 of ECM rigidity is even more pronounced in the chronic colitis model than in severe acute disease. This softening corresponds with the increasing structural deterioration and perforation of the matrix, as revealed by imaging.

Matrisome compositional dynamics reflects the inflammatory state of the colon

To enhance the definition of the two distinct silent pre-symptomatic states, for our third
255 dimension of Multi-ECM, we next addressed the proteomic composition of the ECM – the *matrisome* [51]. We used mass-spectrometry (MS)-based analysis of the tissue proteome, followed by statistical analysis and computational modeling to analyze tissue samples in each experimental state. Eight to ten colon samples were analyzed from in each of the five states: healthy WT, the two pre-symptomatic states (i.e., day 4 DSS, healthy IL-10), and two types of
260 intestinal inflammation (acutely ill day 10 of DSS, chronically ill IL-10). Out of 2045 total proteins identified in all states, 110 ECM and ECM-related proteins, comprising 5% of all identified proteins, were retrieved. Annotation as an ECM protein was made according to the Matrisome, an ECM protein cohort defined by Naba *et al* [51].

To identify the ECM proteins that are significantly changed in abundance during illness, we
265 performed statistical analysis, comparing pairs of states, using a t-test with a Bonferroni correction for multiple (6) pairwise comparisons (corrected $\alpha=0.05/6=0.0083$), and a 0.05 permutation-based false discovery rate (FDR) cutoff. A scheme summarizing our results is

shown in **Fig. 4A**, and in more detail in **Table S3**. As expected, we found that the ECM composition of the healthy WT state is very different from that of both states of illness. Perhaps less expectedly, we also detected a statistically significant proteomic compositional shift in both pre-symptomatic states compared to the healthy WT. A large portion of this protein abundance shift between WT and pre-symptomatic states (i.e., increase in Collagen alpha-2(I) chain, Fibrillin-1 and Transglutaminase-3; decrease in Mucin-2 and lectin mannose-binding-1) is shared between the two pre-symptomatic states (**Fig. 4B**). Remarkably, these two states exhibit the highest compositional similarity of any pair, with no differentially abundant ECM proteins at all (**Fig. 4A**). Moreover, the ECM matrisomic shift from each pre-symptomatic state to its corresponding ill state also shows a shared signature (**Fig. 4C**). The two ill states, though clinically and functionally distinct, are also quite similar in terms of their matrisome, as summarized in **Fig. 4A**. However, using simple statistical analysis, the composition of the chronically ill IL-10 ECM appears to be the most different from its pre-symptomatic counterpart, healthy IL-10 (**Table S3**), which is consistent with our finding that its rigidity is also the lowest of all states.

In summary, MS proteomic analysis of all experimental samples provides a molecular signature for the "pre-symptomatic ECM scaffold" in both acute and chronically ill models, which is compositionally distinct, in its ECM from healthy WT or ill states, but is similar between inflammation models.

Machine-learning-based analyses of ECM proteomics data decode composition and highlight potential biomarkers

In order to further decode our proteomic analysis, we have utilized machine-learning algorithms. Initially, we subjected the data to a supervised learning algorithm, a J48

pruned decision tree classifier, so that we could obtain a visual model to permit nomination of prominent proteins for distinction of clinical states. In our data, identification of such distinguishing features would depend upon correct automated classification of samples into either "healthy WT", "pre-symptomatic" or "ill" classes. As input for this classifier, we used the ECM proteomic data (i.e., 110 matrisome proteins) described above, comprising 8 healthy WT samples, 16 pre-symptomatic samples (day 4 of DSS and healthy IL-10) and 20 ill samples (day 10 of DSS and ill IL-10), where each sample came from a different mouse. We selected a 10-fold stratified cross-validation for the classifier. The resulting model is illustrated in **Fig. 5A**. The algorithm “chose” two proteins, Collagen XVIII/Endostatin (Col18a1) and Fibrillin-1 (Fbn1), as the nodes of the tree. Col18a1 is a ubiquitous non-fibrillar collagen, and a crucial, structurally complex basement membrane proteoglycan[52,53]. The glycoprotein Fibrillin-1 (Fbn1), is part of the microfibril-forming Fibrillin family that is important for connective tissue integrity and function[54]; notably, mutations in the FBN1 gene lead to Marfan syndrome, which mostly manifests in skeletal, cardiovascular, and ocular abnormalities [55,56]. As distinguishing characters of our samples, high protein levels of Col18a1 is a feature of samples from ill mice. Samples with low Col18a1 are further evaluated for a dependent classifier, Fbn1 levels; here, samples with low Col18a1 and high levels of Fbn1 are classified as pre-symptomatic, while samples with low Col18a1 and low levels of Fbn1 are classified as healthy colons (**Fig. 5A**). Plotting all samples according to the abundances of these two proteins resulted in a clear separation between the healthy, pre-symptomatic and ill states (**Fig. 5B**). This analysis allows automated separation of different clinical states from existing large data sets. In order to understand the etiology of this differential protein abundance, we used real-time qPCR to compare transcript levels for these two genes in samples from each disease model. We detected

a significant increase in Col18a1 expression in both ill states (**Fig. 5C**), while Fbn1 expression
315 does not change significantly. This indicates that upregulation of Col18a1 occurs at the transcriptional level, while Fbn1 is most likely regulated at later stages of expression (or by differential degradation), a more common regulatory pathway for ECM proteins that must be secreted to extracellular space and then assembled/modified. Remarkably, these observations of protein and transcript abundance exhibit similar patterns of relative abundance in clinically
320 isolated human biopsies, as described below.

To bridge the gap between the 'omics data and tissue morphology and function, we further analyzed the data holistically, using unsupervised learning tools. As 110 ECM protein levels is a relatively large number of dimensions for describing 44 samples, we used clustering to define groups of proteins that vary similarly across the experimental conditions. This allowed us
325 to identify the dominant patterns in our high-dimensional dataset and avoid overfitting. We identified three clusters of proteins with similar abundance patterns (**Fig. 5D**). We then used principle component analysis (PCA) on the means of each cluster to find the axes of maximal variation. We found that the first PC (PC1) accounted for most (86%) of the variance. PC1 is dominated by proteins in cluster 1 (~75%) and therefore this cluster explains most of the
330 variance between samples. That this analysis is “unsupervised” means that the machine-learning algorithm that "decides" how to compute the PC1 score is only given the ECM composition of each sample, but is not given any other information about each sample; crucially, it has no access to information on which of the five states corresponds to each sample (or the amount of structural damage it displays). Interestingly, the proteins in cluster 1, which are mostly protease
335 inhibitors (e.g., serpins) and coagulation factors (e.g., subunits of the fibrinogen complex) are more abundant in the ill states; while the proteins in cluster 2, which are mostly structural ECM

proteins (e.g., proteoglycans, glycoproteins and collagens), which explain the remainder of sample variance, are reduced in the ill states (**Table S4**). Strikingly, using PC1, the samples were ranked from the healthy WT to similarly ranked pre-symptomatic states (first healthy IL-10 and then day 4 DSS), followed by chronically ill IL-10 and finally, the acutely ill DSS day 10 state. Thus differentiation of samples mostly according to expression of proteins in cluster 1 is sufficient to correctly distinguish between the ill and healthy states (**Fig. 5E)**. The two states at the edges of this ECM ranking, or hierarchy, represent the two extremes of their clinical states. Since values of PC1 approaching the values of healthy WT distribution represent a stable, healthy state and values around day 10 represent illness and instability, we note that healthy IL-10 (as a chronic pre-symptomatic state) is closer to the "healthy edge" than the transient pre-symptomatic state of day 4 of DSS. This observation is consistent with our ECM rigidity analysis, described above. PC1 values of chronic illness reflect more stable disease than the acute one, and are therefore not as close to the "ill edge", as our model would predict. Most importantly, the compositional healthy-ill ranking corresponds to the degree of structural damage that we observe via imaging.

Extracellular matrix proteolysis reveals its role as a metabolite reservoir that is modified under different inflammatory conditions

ECM degradation by proteolytic enzymes, such as matrix metalloproteinases (MMPs), takes place both in homeostasis and during inflammation, predominantly affects its structure. But this proteolysis is also thought to play a more a direct role in signaling, by releasing growth factors and cytokines enmeshed within [57–61], making the ECM a reservoir of signaling intermediates. However, little experimental data exists that directly address the capacity of ECM to function as a metabolite reservoir. We sought to directly test whether ECM degradation leads

360 to release of metabolites, and how the repertoire of metabolites changes according to the unique ECM substrate of each specific tissue state. This metabolomics analysis comprises the fourth dimension of Multi-ECM.

For this experiment, we used decellularized colons from each of the five health-disease states from our mouse studies, and subjected them to ECM degradation by the collagenase 365 MMP-1, followed by global NMR-based metabolomics. As a first step, metabolomic analysis comparing the assay solution I which the ECM was immersed without MMP-1 to assay solution supplemented with MMP-1 showed that without the enzyme, the total metabolite concentration was very low, and the discrimination accuracy among samples was low accordingly (**Fig. S3**). In contrast, MMP-1 treatment in this assay resulted in release of metabolites, mainly amino acids, 370 into the assay solution (**Table S6**). Furthermore, the ECM-derived metabolites from each clinical state released a different combination of metabolites, with a PCA-CA model providing 68% discrimination accuracy among samples (**Fig. 6A**). Univariate analysis of specific metabolites exhibiting significant differential release among all states reveals an emergent pattern: fewer amino acid metabolites are released from pre-symptomatic states compared to both ill states and 375 healthy WT ECMs. However, hydroxy-L-proline, which is a collagen-unique amino acid, is especially enriched among released metabolites from day 4 ECM (**Fig. 6B**). Consistent with our structural and proteomic analyses, pre-symptomatic states are very similar to one another in their ECM metabolite profiles, as revealed by a PCA-CA prediction model in which day 4 samples were used as the test set (**Fig. 6C**).

380 These results suggest that the ECM can serve as a metabolite reservoir in the tissue, releasing these molecules as a result of proteolysis. As proteolysis occurs mainly in inflammation, this reservoir is exhausted to a greater extent. Most importantly, the pre-

symptomatic states release a distinct profile of metabolites, adding another parameter to their unique ECM signature.

385 ***Human IBD recapitulates murine matrisomic dynamics and compositional diversity.***

In order to assess to clinical relevance of our findings in mouse models to human disease, we first tested whether the murine ECM proteomic analysis could be confirmed using an independent human dataset. One existing proteomic dataset comparing healthy donor biopsies with those of UC patients, deposited in the ProteomeXchange, was used for this experiment [62].
390 We examined the relative abundance levels of the proteins that were significantly differentially expressed between the healthy mice (both WT and IL-10) and their corresponding fully diseased sample types (**Fig. 7A**). Remarkably, we found that a significant proportion of these proteins exhibit the same trends in relative abundance in the human patient samples: proteins that are more abundant in ill mice than healthy mice are also more abundant in UC patients than in
395 healthy donors, and vice versa (**Fig. 7B**).

Next, we examined expression of these same putative biomarkers in a new set of human biopsies. For this, we compared mRNA expression levels of Collagen XVIII and Fibrillin-1 between healthy donor biopsies and tissues samples from IBD patients that were endoscopically and histologically identified as inflamed (sample list in **Table S7**). We found that both Collagen
400 XVIII and Fibrillin-1 are significantly upregulated, in concordance with the decision tree applied to the proteomic data from our murine colitis samples (**Fig. 7C**). To corroborate this finding on the protein level and tissue localization, we also stained for Fibrillin-1 in inflamed rectal biopsies of pediatric UC patients with distal colitis. We have found that a subset (~50%) of these patients display a strong epithelial staining of Fibrillin-1 ("FBN1 positive"), while the rest look like the
405 control with little to no staining (**Fig. 7D-E**).

Taken together, analysis of published human proteomic datasets, qPCR and immunostaining of clinical biopsies, validate our murine findings. We confirmed the predictive value of biomarkers Collagen XVIII and Fibrillin-1, which we identified using a machine learning classification analysis, by showing that their expression levels differ significantly in IBD patients compared to healthy donors, and that a subset of UC patients are FBN1 positive at the protein level. Proteomics with machine learning classification therefore represents a powerful application of one of the components of Multi-ECM. We propose this series of techniques, which we collectively call Multi-ECM, as a novel approach to the identification of further biomarkers and clinical diagnostic features in inflammatory diseases, using advanced interrogation of ECM dynamics.

Discussion

In this paper, we have presented Multi-ECM, an integrated, multidisciplinary molecular analysis that offers a sensitive, clinically relevant classification of ECM remodeling. We have used Multi-ECM to improve the evaluation of the heterogeneous, dynamic and hierarchical nature of the ECM as an integral tissue component. As a test case for this method, we used intestinal inflammation, which is both extensively studied clinically and has several well-studied mouse models, and applied the four dimensions of our method to their differentiation. The resulting data streams and analysis provide a high definition representation of healthy intestinal ECM and its remodeling during either acute or chronic inflammation, and most strikingly, uncovered a silent pathology detectable even before the onset of clinical symptoms (**Scheme 1**). Our comprehensive analysis yielded a number of interesting observations with diagnostic and mechanistic implications.

First, our analysis uncovered what we term a "pre-symptomatic ECM scaffold", a reproducible molecular signature that is virtually identical in structure, composition, and ECM-derived metabolite repertoire in acute and chronic models, and crucially, is distinct from that of the healthy WT mouse. Multi-ECM exposed several discrete layers and aspects of ECM remodeling that permitted this discovery. This involves collagen degradation and differential proteolytic activity of gelatinases, supporting MMP-9 as a therapeutic target in IBD[31]. Another strong indication of collagen-degradation processes in distinguishing these states is the prominent accumulation of the collagen-unique amino acid, hydroxyproline[63], which is released by proteolysis, especially at day 4. This may indicate that collagen has begun degrading at this early state, "trapping" hydroxyproline within the matrix. We propose that this proteolyzed pre-symptomatic ECM structure, which SEM shows is highly perforated, and importantly, the mesh underlying the epithelium, allows easier access for immune cell infiltration. Our findings add significant molecular definition to the inflammatory process that precedes appearance of clinical symptoms, by adding to previous work which defined only a single early ECM remodeling feature – changes in the deposition of the abundant ECM glycosaminoglycan component, hyaluronic acid[64] – and that, only in the DSS model. Though an active role for the pro-inflammatory ECM niche in driving colitis development remains to be proved, our analysis provides a more comprehensive picture of multidimensional ECM dynamics at colitis-prone states.

The second discrete feature highlighted by Multi-ECM arises from supervised learning analysis on the matrisome, which suggests a convergence of the five states examined into 1) healthy WT, 2) pre-symptomatic, and 3) ill. We identified Collagen XVIII/Endostatin and Fibrillin-1 as features distinguishing these states that can serve as potential biomarkers for

disease state stratification – indicating either ill or pre-symptomatic states, respectively. We found that both genes are transcriptionally upregulated in human IBD. Neither of these proteins has been extensively studied in the context of IBD previously. Recently, a single report showed that upregulation of Fibrillin-1 gene expression at diagnosis of pediatric IBD predicts the fibrostenotic complication of the disease[65]. This suggests the subset of patients with FBN1-positive epithelial staining in our study as candidates for long term clinical follow-up, to determine whether this molecular feature may have similar prognostic impact. Thus, machine learning as part of Multi-ECM has already demonstrated its potential for identifying relevant clinical markers and disease prognostic indicators.

A third advance of Multi-ECM arising from unsupervised analysis of ECM compositions, is the independent stratification of the five states by reducing proteomic composition to one principle component. Remarkably, the PC1 score output by this unsupervised analysis nicely differentiates between the five states, and moreover, orders the five states in complete correspondence to molecular details uncovered by other components of Multi-ECM. Principal component analysis produced a ranking in perfect concordance with the amount of structural damage in each state, from the dense and well-structured ECM of the healthy WT tissue to the bluntly damaged ECM of acute inflammation, which includes complete erosion of crypt macrostructure. For example, the IL-10 mouse exhibits somewhat lower structural damage than the WT mouse state in the same group, and correspondingly the PC1 score is higher. Hence, our comprehensive analysis addresses the fundamental need for a method for mapping of higher order structural organization to protein amounts, measured by quantitative methods.

Finally, synthesizing the data generated by Multi-ECM, we can also show how the phenotypic distinction between acute and chronic disease is reflected in our analysis and not only

the gross division into three groups that we see at first glance. In our model, the duration and difference in tissue resilience of each disease type explains the position of each of our five states along the heathy-ill axis of PC1, in concordance with microscopic/structural analysis. Starting at the "ill" end of the axis – acute inflammation is more aggressive, but it is also transient and spatially discontinuous, and therefore, it is plausible that under these circumstances, the tissue as a whole can withstand harsh structural damage to the ECM without signaling for broad changes in protein production or extensively affecting tissue rigidity. However, the organized, yet perforated, structure of the colonic ECM in steady-state chronic inflammation the process of colitis management by the tissue to maximize tissue functionality in light of ongoing challenge. This is suggested by the somewhat regular, though altered, crypt pattern. Finally, long-term chronic inflammation would be predicted to allow more time for ECM reorganization, including broader compositional and mechanical changes that support this tissue resilience and perhaps promote maintenance of inflammation. Unexpectedly, our model can also account for observed, subtle differences between the two pre-symptomatic states; the rigidity of the healthy IL-10 more closely resembles its WT counterpart than the other pre-symptomatic state, which is reflected by its position on the PC1 ranking. We posit that since IL-10 mice can be healthy for long periods of time, in pathogen-free conditions, until perturbation by piroxicam, they exhibit a “healthy” mechanical state. In contrast, day 4 of the acute model is inherently a transient state, and therefore already displays the initial mechanical softening and compositional tendency (PC1 score) associated with colitis. This model, describing the response of gut tissue to acute and chronic disease, both in ill and pre-symptomatic states, arises from our multi-pronged data set, and includes both molecular and tissue-level phenotypes.

In summary, our integrative approach called "Multi-ECM" permits a well-rounded picture, encompassing multiple levels of detail, of the ECM – an enigmatic, yet substantial, tissue component. We conclude that by manifesting the pathophysiological state of a tissue, the ECM reveals silent disease-prone states that occur before clinical symptoms are detected. By applying this integrative analysis to the case of colitis, we found a general pre-symptomatic state of intestinal inflammation that can be either brief or long-lasting. Within the realm of IBD, Multi-ECM has the potential to unravel the early molecular events leading up to the tissue-damaging complications of IBD, fibrostenosis and fistulization. Understanding how early ECM remodeling events leave the tissue vulnerable or prone to pathology suggests an avenue for development of novel therapeutic strategies. Our results suggest that the ECM is a dynamic and active player in tissue health that transmits a variety of biological signals and resources, and is not merely a passive structural entity. Adoption of Multi-ECM as a component of future clinical and mechanistic studies of inflammation, as well as other pathologies, should only add to this empowering narrative.

Materials and Methods:

Animals

Seven-week old C57BL/6 male mice were purchased from Envigo and were allowed to adapt for one week before experimental procedure. IL-10$^{-/-}$ C57BL/6 mice from Jackson
515 Laboratories were inbred at the Weizmann Institute of Science, and experiments were performed on six to eight-week old male mice. Healthy naïve IL-10$^{-/-}$ were eight to twelve weeks old at time of harvest. All experiments and procedures were approved by the Weizmann Institute of Science animal care and use committees, protocol numbers 02230413-2, 06481012-1.

Human samples

520 Biopsies were collected from healthy donors and from mucosa of IBD patients after providing informed consent in accordance with the ethical standards on Human Experimentation and the Declaration of Helsinki (protocol # 0725-14-TLV and # 3312-16-SMC). Samples were classified as being inflamed or normal by endoscopy and histology. Surgically resected inflamed colon samples from IBD patients were obtained from the Israel National Biobank (MIDGAM).

525 Samples were taken from both male and female subjects. The samples in each group for the RNA analysis are listed in **Table S7**. Samples used in RNA analysis were snap-frozen in liquid nitrogen until RNA extraction. Biopsies used for immunohistochemical stains were fixed in 4% paraformaldehyde.

Intestinal inflammation induction and evaluation

530 Acute colonic inflammation was induced by administration of 1.25% DSS (MP Biomedicals LLC) in drinking water of C57BL/6 mice for seven days[38]. Chronic colonic inflammation was accelerated and synchronized by peroral administration of 200ppm piroxicam (Sigma-Aldrich ltd.) to IL-10$^{-/-}$ C57BL/6 mice via supplementation in normal murine chow[43]. Mice were

weighed 3 times a week over the course of the experiment. Colitis progression was evaluated over the course of the experiment using the Karl Stortz Coloview mini endoscope system and colonic inflammation was scored (0-15) as previously described[40]. Inflammation scores were categorized according to the following: healthy (0-4); mildly inflamed (5-7); inflamed (8-11); severely inflamed (12-15). Another form of inflammation evaluation was histological analysis by H&E staining of formalin-fixed paraffin-embedded tissues sections. The degree of histological damage and inflammation was graded in a blinded fashion by an expert pathologist. The following parameters were scored in the DSS-induced colitis model: amount of inflammation (0, none; 1, mild; 2, moderate; 3, severe; 4, accumulation of inflammatory cells in the gut lumen); percentage of colon area involved (0, none; 1, <=10%; 2, 10%-30%; 3, 30%-50%; 4, 50%-90%; 5, >90%); depth of inflammation and layers involved (0, none; 1, mucosa only; 2, mucosa and submucosa; 3, limited transmural involvement; 4, transmural); ulceration (0, <30%; +1, >30%; -1, regeneration). The overall histological score was the sum of the three parameters (maximum score of 14). Histopathological scoring for the IL-10 model was performed in a similar fashion, but only considering the amount of inflammation and percentage of colon area involved (maximum score of 9).

Decellularization of colonic tissue

Decellularization was performed as previously described[66]. Samples were frozen in ddH_2O and subjected to six freeze-thaw cycles with replacement of ddH_2O each time. Subsequently, samples were immersed in 25mM NH_4OH (Merck) for 20 minutes. Finally, isolated ECM scaffolds were washed six times in ddH_2O and stored at 4˚C until use.

555 Second-harmonic generation (SHG) microscopy

Native snap-frozen murine colon samples were thawed in PBS and imaged using a two-photon microscope (2PM:Zeiss LSM 510 META NLO; equipped with a broadband Mai Tai-HP-femtosecond single box tunable Ti-sapphire oscillator, with automated broadband wavelength tuning 700–1,020 nm from Spectraphysics, for two-photon excitation). For second-harmonic
560 imaging of collagen, a wavelength of 800-820 nm was used (detection at 390-450nm). Quantification of crypt diameter was done using ImageJ software (Research Service Branch, NIH).

Scanning electron microscope (SEM)

Decellularized mouse colon tissues were used and prepared as described in [67]. Then,
565 samples were subjected to critical point dehydration (CPD) and attached to a carbon sticker. Finally, samples were coated with gold palladium, and imaged under SEM Ultra (Carl Zeiss international).

In situ zymography

In situ zymography was conducted as previously described[47] on unfixed 10µm
570 transverse mouse colon sections using DQ Gelatin™ (Invitrogen) from pig skin. Following, sections were fixed with 4% paraformaldehyde, and cell nuclei were stained with DRAQ5 (Thermo Scientific). Finally, mounting solution (Immu-Mount™ Thermo Scientific) was added and slides were covered with a coverslip. The slides were imaged under a Leica TCS SP8 microscope (both confocal and multiphoton). DQ gelatin was excitated at 488nm and its
575 emission was detected at 515nm.

Testing elastic properties of ECM samples by Atomic Force Microscopy

The AFM analysis was carried out on rehydrated ECM samples derived from mouse colon tissues. Typically, three-four samples were analyzed for each inflammatory condition. The rehydration process was as follows: ECM samples were grossly dried and attached to glass
580 coverslips (diameter 15 mm) by means of a thin bi-adhesive tape. Samples were then attached to the bottom of Petri dishes (Greiner Bio-One) and left overnight in an evacuated desiccator in order to dry out, so to improve spreading and adhesion of ECM on tape. Prior to AFM measurements, the Petri dish hosting the ECM sample was filled with ddH_2O and the ECM was allowed to rehydrate for 30 minutes at room temperature. Measurements were carried out at
585 room temperature.

To accomplish the nanomechanical characterization of the ECM samples, a Bioscope Catalyst AFM (Bruker) was used to collect series of force vs distance curves[68,69], recording the local deformation of the sample induced by the AFM probe. According to a recently published protocol[50], we used monolithic borosilicate glass probes consisting in micrometer-
590 sized spherical glass beads with radius R=9-10 μm attached to silicon cantilevers with a force constant k=0.25-0.35 N/m. The probes were produced and characterized as previously described[49].

Each set of force curves (a force volume) consisted of a 16x16 array of curves acquired on a 70μm × 70μm area. Ten force volumes were typically recorded on each ECM on
595 macroscopically separated regions. All measurements were performed with the following parameters: 4096 points per curve, ramp length $L=10\ \mu m$, maximum applied force $F = 60\text{-}70nN$, and ramp frequency $f=1.10\ Hz$. Typically, indentations up to 2-3 μm were obtained. Data processing of force volumes was carried out in Matlab according to the general protocol described previously[50]. The values of the Young's Modulus (YM) were extracted by fitting the

Hertz model to the experimental data[68,69]. A first very soft indentation region (0-40% of total indentation) was excluded, in order to separate the possible contribution of loosely-bound superficial layers from those of the bulk ECM. Artifacts derived from an ill-defined contact area between sample and probe, like boundaries of colonic crypts or more generally crypts with characteristic dimensions comparable to, or larger than the probe diameter, were identified and discarded.

The distributions of YM values of the ECMs from mice turned out to be the envelope of several nearly lognormal modes, representing the major contributions to the overall ECM rigidity and originating from micro-scale domains that the AFM probe was able to resolve. Each mode in the distributions represents the elastic properties of a region of the ECM with lateral dimensions (and thickness) of several microns, which may define a structural and functional domain of the ECM. Multi-Gaussian fit in semilog10 scale allowed identification of the peak value E' and the geometric standard deviation σ_g^{10} of each lognormal mode; from these values the median value E_{med} and the standard deviation of the median σ_{med} were evaluated for all modes as $E_{med} = 10^{E'}$ and $\sigma_{med} = \sqrt{\pi/2} E_{med} \sigma_g^{10} / \sqrt{N}$ [70], N being the number of force curves in each mode. The effective rigidity of each ECM sample was characterized by the weighted average of median values $E = \sum_i f_i E_{med,i}$, using the fraction $f_i = N_i / N_{tot}$ of force curves in the mode as the weight; the total error σ_E associated to E was calculated by summing in quadrature the propagated error of the medians $\sigma = \sqrt{\sum_i f_i^2 \sigma_{med,i}^2}$ and an effective instrumental relative error $\sigma_{instr} = 3\ \%$: $\sigma_E = \sqrt{\sigma_{instrum}^2 E^2 + \sigma^2}$. The average median values of the YM of the different states of inflammation have also been evaluated; the corresponding error has been calculated as the standard deviation of the mean summed in quadrature with the propagated σ_E.

ECM proteomics by liquid chromatography–tandem mass spectrometry (LC-MS/MS) analysis

Tissue slices from colons of WT or IL-10 mice at different stages of the inflammation models were immersed in a solution containing 50% trifluoroethanol (TFE), 25mM ammonium bicarbonate (Sigma Aldrich) and 15mM dithiothreitol (DTT) and coarsely homogenized by repeated cycles of boiling, freezing and sonication. Subsequently, samples were shaken at 30˚C, 1400 rpm for 30 minutes, followed by the addition of iodoacetamide (Sigma Aldrich) to a final concentration of 25mM and further shaking for 30 minutes at 30˚C and 1400 rpm. TFE was then diluted to 25% with 50mM ABC, LysC (Wako Chemicals, 1:100 lysC:protein ratio) and sequencing-grade modified trypsin (Promega, 1:50 trypsin:protein ratio) was added and incubated overnight at room temperature. On the following morning, more trypsin was added (1:80 trypsin: protein ratio) for 4 hours. Peptide mixtures were purified on C18 stage tips. Eluted peptides were loaded onto a 50 cm long EASY-spray reverse phase column and analyzed on an EASY- nLC- 1000 HPLC system (Thermo Scientific) coupled to a Q-Exactive Plus MS (Thermo Scientific). Peptides were separated over 240 minutes with a gradient of 5−28 % buffer B (80% acetonitrile and 0.1% formic acid). One full MS scan was acquired at a resolution of 70,000 and each full scan was followed by the selection of 10 most intense ions (Data dependent Top 10 method) at a resolution of 17,500 for MS2 fragmentation.

Raw MS data was analyzed with MaxQuant software[71,72] (version 1.5.2.18) with the built-in Andromeda search engine[73] by searching against the mouse reference proteome (UniprotKB,Nov2014). Enzyme specificity was set to trypsin cleavage after lysine and arginine and up to two miscleavages were allowed. The minimum peptide length was set to seven amino acids. Acetylation of protein N termini, deamidation of asparagine and glutamine, and oxidation of methionine were set as variable modifications. Carbamidomethylation of cysteine was set as a

645 fixed modification. Protein identifications were sorted using a target-decoy approach at a false discovery rate (FDR) of 1% at the peptide and protein levels. Relative, label-free quantification of proteins was performed using the MaxLFQ algorithm integrated into MaxQuant environment with minimum ratio count of two[74].

Computational analysis of MS data

650 Classification was done using a J48 decision tree algorithm[75] with a stratified 10-fold cross-validation in the Weka software for machine learning (University of Waikato). Clustering according to squared Euclidean distance and principal component analysis (PCA) were performed in Matematica (Wolfram Research).

Quantitative Real Time PCR on mouse and human colon tissues

655 Colon tissues were homogenized using a bead beater homogenizer. Total RNA from was isolated using the PerfectPure RNA Tissue Kit (5 Prime GmbH) according to the manufacturer's protocol. 0.6-1 µg of total RNA is reverse transcribed using High Capacity cDNA Kit (Applied Biosystems inc.). qRT-PCR was performed using SYBR Green PCR Master Mix (Applied Biosystems inc.) on an ABI 7300 instrument (Applied Biosystems). Values are normalized to an

660 Actin-β control in mouse tissues and to Hypoxanthine-guanine Phosphoribosyltransferase (HPRT) and Ubiquitin C (UBC) controls in the human tissues. Primer sequences are listed in **Table S5**. Data is presented as mean fold change compared to either healthy WT in mouse tissues or to normal, non-inflamed, human tissues using the $2^{-\Delta\Delta CT}$ method[76]. Mice samples were grouped according to general state: (i) healthy WT, (ii) pre-symptomatic (day 4 of DSS and

665 healthy IL-10) and (iii) ill (day 10 of DSS and ill IL-10). Standard error of the mean (s.e.m) was calculated on the $2^{-\Delta CT}$ data, as was the statistical analysis.

Extraction of ECM metabolites

Decellularized colons were immersed in TNC buffer (50mM Tris pH 7.5, 200mM NaCl and 5mM $CaCl_2$) and incubated with 50nM human recombinant matrix metalloproteinase (MMP)-1 for 24 hours in 37°C. The solution of degradation products was taken for NMR analysis.

NMR-based Metabolomics

Samples of frozen extracts were thawed at room temperature and shaken before use. A 630 µL aliquot of each sample was added to 70 µL of potassium phosphate buffer (1.5 M K2HPO4, 100% (v/v) 2H2O, 10 mM sodium trimethylsilyl [2,2,3,3-2H4]propionate (TMSP), at pH 7.4). 600 µL of each mixture were transferred into 4.25 mm NMR tubes (Bruker BioSpin srl) for analysis.

^{1}H NMR spectra were acquired using a Bruker 600 MHz metabolic profiler (Bruker BioSpin) operating at 600.13 MHz proton Larmor frequency and equipped with a 5mm TXI 1H-13C-15N and 2H-decoupling probe including a z axis gradient coil, an automatic tuning-matching (ATM) and an automatic sample changer (SampleJet). A BTO 2000 thermocouple served for temperature stabilization at the level of approximately 0.1 K at the sample. Before measurement, samples were kept for at least 5 minutes inside the NMR probehead for temperature equilibration (300 K). For each sample, a 1H NMR spectrum was acquired with the Carr, Purcell, Meiboom, and Gill (CPMG) sequence using a 1D spin–echo sequence with water presaturation. 256 scans over a spectral region of 12 kHz were collected into 64 K points, giving an acquisition time of 3.06 s .

The raw data were multiplied by a 0.3 Hz exponential line broadening before Fourier transformation into 128 K points. Chemical shift was referenced to the signal of TMSP at δ 0.00

690 ppm. Bucketing was performed to the data dividing each spectrum in the region 10.0 - 0.2 ppm into sequential segments ("bins") of 0.02 ppm width which were integrated using AMIX software (Bruker BioSpin). The spectral region 6.0 - 4.2 ppm, containing the water signals, was discarded. Finally, total area normalization was carried out.

Immunohistochemical (IHC) stain for Fibrillin-1

695 Tissues were fixed in 4% paraformaldehyde and embedded in paraffin, and serial 5-µm sections were prepared from the whole biopsy for IHC. Sections were de-paraffinized and epitope retrieval was performed. Sections were incubated for 1 h with anti Fibrillin1 (Abcam), diluted 1:100, at RT followed by secondary antibody (HiDef Detection Polymer, 954D, Cell Marque, Rocklin, CA, USA). A peroxidase substrate kit (SK-4100, Vector Labs, USA) was used 700 as a chromogen and hematoxylin as a counterstain. Tissue exposed only to the secondary antibody was used as negative control.

Staining coverage was quantified as the mean percentage of area covered by staining out of the total image in 3-10 fields of view per sample. The analysis was carried out using ImageJ software (Research Service Branch, NIH) by applying color deconvolution for DAB staining, 705 followed by a consistent binary threshold.

Quantification and Statistical Analysis

General

Statistical parameters including the exact value of n, the definition of center, dispersion and precision measures (mean ± SEM) and statistical significance are reported and portrayed in 710 the figures and the figure legends. Wherever outlier removal is mentioned, outliers are considered values that are over 3 median absolute deviations distance from the median, which is a robust outlier detection method[77]. Data is judged to be statistically significant when $P < 0.05$

by a two-tailed student's t test or with a Bonferroni correction of 0.05/m, where m=number of pairwise comparisons. In figures, asterisks denote statistical significance as calculated by student's t test (∗, $P < 0.05$ or 0.05/m; ∗∗, $P < 0.01$ or 0.01/m; ∗∗∗, $P < 0.001$ or 0.001/m) in Microsoft Office Excel.

Statistical analysis of MS data

Bioinformatics analysis was performed with the Perseus program version 1.5.1.4 [78]. The proteomic data was first filtered to remove the potential contaminants, proteins only identified by their modification site and reverse proteins. Next, the intensity values were log2 transformed and data was filtered to have at least three valid values in each group. Missing values were imputed based on normal distribution and differentially abundant proteins were chosen according to a student's t-test with a Bonferroni correction for multiple (6) pairwise comparisons (corrected $\alpha=0.05/6=0.0083$), and a 0.05 permutation-based false discovery rate (FDR) cutoff for analysis of multiple proteins [79]. Statistical significance of intersections in Venn diagrams and proportion of differentially abundant ECM proteins out of all differentially abundant proteins was assessed using a hypergeometric probability density function in Matlab (Matworks inc.).

Statistical analysis of metabolomics

Various kinds of multivariate and univariate statistical techniques were applied on the obtained buckets using R 3.0.2 in house scripts. Principal Component Analysis (PCA) was used to obtain a preliminary outlook of the data (visualization in a reduced space, clusters detection, screening for outliers). Canonical analysis (CA) was used in combination with PCA to increase the supervised separation of the analyzed groups. Accuracy, specificity and sensitivity were estimated according to standard definitions. The global accuracy for classification was assessed

by means of a Montecarlo cross-validation scheme. Several metabolites, whose peaks in the spectra were well defined and resolved, were assigned. Signal identification was achieved using a library of NMR spectra of pure organic compounds, public databases (such as HMBD, Human Metabolic Database) storing reference NMR spectra of metabolites, spiking NMR experiments and literature data. The relative concentrations of the various metabolites in the different spectra were calculated by integrating the signal area. One-way ANOVA was used for the determination of the meaningful metabolites; a p-value ≤ 0.05 was considered statistically significant. The changes in metabolites levels between two groups of spectra are calculated as the log2 Fold Change (FC) ratio of the normalized median intensities of the corresponding signals in the spectra of the two groups.

Acknowledgments

We thank Dr. Ori Brenner for assistance in histopathological scoring of intestinal inflammation. Also, we thank Dr. Inna Solomonov for discussions and insights on data presentation and analysis. Finally, we thank Dr. Miriam Rosenberg for helpful comments on the text and styling.

Funding: IS is the Incumbent of the Maurizio Pontecorvo Professorial Chair and has received funding from the Israeli Science Foundation (1226/13), the European Research Council AdG (THZCALORIMETRY - DLV-695437) and the USA-Israel Binational Science Foundation (712506-01), the Ambach fund, and the Kimmelman center at the WIS. AP thanks the Dept. of Physics of the University of Milano for financial support under the project "Piano di Sviluppo dell'Ateneo per la Ricerca 2014", actions: "A. Upgrade of instrumentation", and "B. Support to young researchers". AS and TG are funded by the Israel Ministry of Science Technology and Space.

Author contributions

Conceptualization, ES and IS; Methodology, ES, RA, AS, MA, LP and VG; Formal Analysis, ES, AS, MA, LP and VG; Investigation, ES and RA; Resources, OM, NG, SF, LW, DSS, CV, AP, TG, PM; Writing – Original Draft, ES; Writing – Review & Editing, ES, IS; Visualization, ES, MA, LP, VG and AP; Supervision, CV, DSS, AP, PT, TG, PM, CL, UA and IS ; Funding Acquisition, IS, AP, TG, PM.

Data and materials availability

The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD004740 and can be viewed by logging in with the following details: Username: reviewer98510@ebi.ac.uk; Password: hKx9zFSu.

References:

1. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. J Cell Sci. 2010;123: 4195–200. doi:10.1242/jcs.023820
2. Bosman FT, Stamenkovic I. Functional structure and composition of the extracellular matrix. J Pathol. 2003;200: 423–8. doi:10.1002/path.1437
3. Hynes RO. The extracellular matrix: not just pretty fibrils. Science. 2009;326: 1216–1219. doi:10.1126/science.1176009
4. Mott JD, Werb Z. Regulation of matrix biology by matrix metalloproteinases. Curr Opin Cell Biol. 2004;16: 558–564. doi:10.1016/j.ceb.2004.07.010
5. Shimshoni E, Yablecovitch D, Baram L, Dotan I, Sagi I. ECM remodelling in IBD: innocent bystander or partner in crime? The emerging role of extracellular molecular events in sustaining intestinal inflammation. Gut. BMJ Publishing Group Ltd and British Society of Gastroenterology; 2015;64: 367–72. doi:10.1136/gutjnl-2014-308048
6. Frey H, Schroeder N, Manon-Jensen T, Iozzo R V, Schaefer L. Biological interplay between proteoglycans and their innate immune receptors in inflammation. FEBS J. 2013;280: 2165–2179. doi:10.1111/febs.12145
7. Gao Q, Meijer MJW, Kubben FJGM, Sier CFM, Kruidenier L, van Duijn W, et al. Expression of matrix metalloproteinases-2 and -9 in intestinal tissue of patients with inflammatory bowel diseases. Dig Liver Dis. 2005;37: 584–92. doi:10.1016/j.dld.2005.02.011
8. Wu F, Vij N, Roberts L, Lopez-Briones S, Joyce S, Chakravarti S. A novel role of the lumican core protein in bacterial lipopolysaccharide-induced innate immune response. J Biol Chem. 2007;282: 26409–17. doi:10.1074/jbc.M702402200
9. Opdenakker G, Van den Steen PE, Van Damme J. Gelatinase B: a tuner and amplifier of immune functions. Trends Immunol. 2001;22: 571–579. doi:10.1016/S1471-4906(01)02023-3
10. Sorokin L. The impact of the extracellular matrix on inflammation. Nat Rev Immunol. Nature Publishing Group; 2010;10: 712–723. doi:10.1038/nri2852
11. de la Motte CA. Hyaluronan in intestinal homeostasis and inflammation: implications for fibrosis. Am J Gastrointest Liver Physiol. 2011;301: 945–949. doi:10.1152/ajpgi.00063.2011.
12. Afik R, Zigmond E, Vugman M, Klepfish M, Shimshoni E, Pasmanik-Chor M, et al. Tumor macrophages are pivotal constructors of tumor collagenous matrix. J Exp Med. 2016; jem.20151193. doi:10.1084/jem.20151193
13. Grossman M, Ben-Chetrit N, Zhuravlev A, Afik R, Bassat E, Solomonov I, et al. Tumor cell invasion can be blocked by modulators of collagen fibril alignment that control assembly of the extracellular matrix. Cancer Res. 2016;
14. Lu P, Weaver VM, Werb Z. The extracellular matrix: A dynamic niche in cancer progression. J Cell Biol. 2012;196: 395–406. doi:10.1083/jcb.201102147
15. Theocharis AD, Skandalis SS, Tzanakakis GN, Karamanos NK. Proteoglycans in health

and disease: novel roles for proteoglycans in malignancy and their pharmacological targeting. FEBS J. 2010;277: 3904–3923. doi:10.1111/j.1742-4658.2010.07800.x

16. Talmi-Frank D, Altboum Z, Solomonov I, Udi Y, Jaitin DA, Klepfish M, et al. Extracellular Matrix Proteolysis by MT1-MMP Contributes to Influenza-Related Tissue Damage and Mortality. Cell Host Microbe. Elsevier; 2016;20: 458–470. doi:10.1016/j.chom.2016.09.005

17. Goldberg R, Meirovitz A, Hirshoren N, Bulvik R, Binder A, Rubinstein AM, et al. Versatile role of heparanase in inflammation. Matrix Biol. 2013;32: 234–40. doi:10.1016/j.matbio.2013.02.008

18. Khokha R, Murthy A, Weiss A. Metalloproteinases and their natural inhibitors in inflammation and immunity. Nat Rev Immunol. 2013;13: 649–665. doi:10.1038/nri3499

19. de Souza HSP, Fiocchi C. Immunopathogenesis of IBD: current state of the art. Nat Rev Gastroenterol Hepatol. Nature Publishing Group; 2016;13: 13–27. doi:10.1038/nrgastro.2015.186

20. Nielsen OH, Rogler G, Hahnloser D, Thomsen OØ. Diagnosis and management of fistulizing Crohn's disease. Nat Clin Pract Gastroenterol Hepatol. 2009;6: 92–106. doi:10.1038/ncpgasthep1340

21. Rieder F, Fiocchi C. Intestinal fibrosis in IBD — a dynamic , multifactorial process. 2009;6: 228–235. doi:10.1038/nrgastro.2009.31

22. Latella G, Rogler G, Bamias G, Breynaert C, Florholmen J, Pellino G, et al. Results of the 4th scientific workshop of the ECCO (I): Pathophysiology of intestinal fibrosis in IBD. J Crohns Colitis. European Crohn's and Colitis Organisation; 2014; doi:10.1016/j.crohns.2014.03.008

23. Kirkegaard T, Hansen A, Bruun E, Brynskov J. Expression and localisation of matrix metalloproteinases and their natural inhibitors in fistulae of patients with Crohn's disease. Gut. 2004;53: 701–709. doi:10.1136/gut.2003.017442

24. Bailey CJ, Hembry RM, Alexander A, Irving MH, Grant ME, Shuttleworth CA. Distribution of the matrix metalloproteinases stromelysin, gelatinases A and B, and collagenase in Crohn's disease and normal intestine. J Clin Pathol. 1994;47: 113–116. doi:10.1136/jcp.47.2.113

25. Koller FL, Dozier EA, Nam KT, Swee M, Birkland TP, Parks WC, et al. Lack of MMP10 exacerbates experimental colitis and promotes development of inflammation-associated colonic dysplasia. Lab Invest. Nature Publishing Group; 2012;92: 1749–59. doi:10.1038/labinvest.2012.141

26. Meijer MJW, Mieremet-Ooms MAC, van der Zon AM, van Duijn W, van Hogezand RA, Sier CFM, et al. Increased mucosal matrix metalloproteinase-1, -2, -3 and -9 activity in patients with inflammatory bowel disease and the relation with Crohn's disease phenotype. Dig Liver Dis. 2007;39: 733–9. doi:10.1016/j.dld.2007.05.010

27. Pedersen G, Saermark T, Kirkegaard T, Brynskov J. Spontaneous and cytokine induced expression and activity of matrix metalloproteinases in human colonic epithelium. Clin Exp Immunol. 2009;155: 257–265. doi:10.1111/j.1365-2249.2008.03836.x

28. Pender SLF, MacDonald TT. Matrix metalloproteinases and the gut - New roles for old enzymes. Current Opinion in Pharmacology. 2004. pp. 546–550. doi:10.1016/j.coph.2004.06.005

29. Baugh MD, Perry MJ, Hollander AP, Davies DR, Cross SS, Lobo AJ, et al. Matrix metalloproteinase levels are elevated in inflammatory bowel disease. Gastroenterology. 1999;117: 814–822.

30. Garg P, Rojas M, Ravi A, Bockbrader K, Epstein S, Vijay-Kumar M, et al. Selective ablation of matrix metalloproteinase-2 exacerbates experimental colitis: contrasting role of gelatinases in the pathogenesis of colitis. J Immunol. 2006;177: 4103–12.

31. Sela-Passwell N, Kikkeri R, Dym O, Rozenberg H, Margalit R, Arad-Yellin R, et al. Antibodies targeting the catalytic zinc complex of activated matrix metalloproteinases show therapeutic potential. Nat Med. Nature Publishing Group; 2012;18: 143–7. doi:10.1038/nm.2582

32. Castaneda FE, Walia B, Vijay-Kumar M, Patel NR, Roser S, Kolachala VL, et al. Targeted deletion of metalloproteinase 9 attenuates experimental colitis in mice: central role of epithelial-derived MMP. Gastroenterology. 2005;129: 1991–2008. doi:10.1053/j.gastro.2005.09.017

33. Salmela MT, Pender SLF, Karjalainen-Lindsberg M-L, Puolakkainen P, Macdonald TT, Saarialho-Kere U. Collagenase-1 (MMP-1), matrilysin-1 (MMP-7), and stromelysin-2 (MMP-10) are expressed by migrating enterocytes during intestinal wound healing. Scand J Gastroenterol. Informa UK Ltd UK; 2004;39: 1095–104. doi:10.1080/00365520410003470

34. Hermano E, Lerner I, Elkin M. Heparanase enzyme in chronic inflammatory bowel disease and colon cancer. Cell Mol Life Sci. 2012;69: 2501–13. doi:10.1007/s00018-012-0930-8

35. Waterman M, Ben-Izhak O, Eliakim R, Groisman G, Vlodavsky I, Ilan N. Heparanase upregulation by colonic epithelium in inflammatory bowel disease. Mod Pathol. 2007;20: 8–14. doi:10.1038/modpathol.3800710

36. Lohr K, Sardana H, Lee S, Wu F, Huso DL, Hamad AR, et al. Extracellular matrix protein lumican regulates inflammation in a mouse model of colitis. Inflamm Bowel Dis. 2012;18: 143–151. doi:10.1002/ibd.21713.Extracellular

37. Stange EF, Travis SPL, Vermeire S, Reinisch W, Geboes K, Barakauskiene A, et al. European evidence-based Consensus on the diagnosis and management of ulcerative colitis: Definitions and diagnosis. J Crohns Colitis. 2008;2: 1–23. doi:10.1016/j.crohns.2007.11.001

38. Okayasu I, Hatakeyama S, Yamada M, Ohkusa T, Inagaki Y, Nakaya R. A novel method in the induction of reliable experimental acute and chronic ulcerative colitis in mice. Gastroenterology. 1990;98: 694–702.

39. Stange EF, Travis SPL, Vermeire S, Beglinger C, Kupcinkas L, Geboes K, et al. European evidence based consensus on the diagnosis and management of Crohn's disease: definitions and diagnosis. Gut. BMJ Publishing Group; 2006;55 Suppl 1: i1-15. doi:10.1136/gut.2005.081950a

40. Becker C, Fantini MC, Neurath MF. High resolution colonoscopy in live mice. Nat Protoc. 2006;1: 2900–2904. doi:10.1038/nprot.2006.446

41. Kühn R, Löhler J, Rennick D, Rajewsky K, Müller W. Interleukin-10-deficient mice develop chronic enterocolitis. Cell. 1993;75: 263–274. doi:10.1016/0092-8674(93)80068-P

42. Glocker E-O, Kotlarz D, Boztug K, Gertz EM, Schäffer AA, Noyan F, et al. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. N Engl J Med. 2009;361: 2033–45. doi:10.1056/NEJMoa0907206

43. Berg D, Zhang J, Weinstock J, Ismail H. Rapid development of colitis in NSAID-treated IL-10–deficient mice. Gastroenterology. 2002; 1527–1542. doi:10.1053/gast.rostaglandins

44. van der Rest M, Garrone R. Collagen family of proteins. FASEB J. 1991;

45. Van den Steen PE, Dubois B, Nelissen I, Rudd PM, Dwek RA, Opdenakker G. Biochemistry and Molecular Biology of Gelatinase B or Matrix Metalloproteinase-9 (MMP-9). Crit Rev Biochem Mol Biol. Taylor & Francis; 2002;37: 375–536. doi:10.1080/10409230290771546

46. Zeng Z-S, Cohen AM, Guillem JG. Loss of basement membrane type IV collagen is associated with increased expression of metalloproteinases 2 and 9 (MMP-2 and MMP-9) during human colorectal tumorigenesis. Carcinogenesis. 1999;20: 749–755. doi:10.1093/carcin/20.5.749

47. George SJ, Johnson JL. In Situ Zymography. Humana Press, Totowa, NJ; 2010. pp. 271–277. doi:10.1007/978-1-60327-299-5_17

48. Kofla-Dlubacz A, Matusiewicz M, Krzystek-Korpacka M, Iwanczak B. Correlation of MMP-3 and MMP-9 with crohn's disease activity in children. Dig Dis Sci. 2012;57: 706–712. doi:10.1007/s10620-011-1936-z

49. Indrieri M, Podestà A, Bongiorno G, Marchesi D, Milani P. Adhesive-free colloidal probes for nanoscale force measurements: production and characterization. Rev Sci Instrum. 2011;82: 023708. doi:10.1063/1.3553499

50. Puricelli L, Galluzzi M, Schulte C, Podestà A, Milani P. Nanomechanical and topographical imaging of living cells by atomic force microscopy with colloidal probes. Rev Sci Instrum. 2015;86: 033705. doi:10.1063/1.4915896

51. Naba A, Clauser KR, Hoersch S, Liu H, Carr SA, Hynes RO. The Matrisome: In Silico Definition and In Vivo Characterization by Proteomics of Normal and Tumor Extracellular Matrices. Molecular & Cellular Proteomics. 2012. p. M111.014647-M111.014647. doi:10.1074/mcp.M111.014647

52. Marneros AG, Olsen BR. Physiological role of collagen XVIII and endostatin. FASEB J. 2005;19: 716–28. doi:10.1096/fj.04-2134rev

53. Utriainen A, Sormunen R, Kettunen M, Carvalhaes LS, Sajanti E, Eklund L, et al. Structurally altered basement membranes and hydrocephalus in a type XVIII collagen deficient mouse line. Hum Mol Genet. Oxford University Press; 2004;13: 2089–2099. doi:10.1093/hmg/ddh213

54. Sakai LY, Keene DR, Engvall E. Fibrillin, a new 350-kD glycoprotein, is a component of extracellular microfibrils. J Cell Biol. 1986;103: 2499–509.

55. Dietz HC, Cutting CR, Pyeritz RE, Maslen CL, Sakai LY, Corson GM, et al. Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. Nature. 1991;352: 337–339. doi:10.1038/352337a0

56. Hollister D, Godfrey M, Sakai L, Pyeritz R. Immunohistologic Abnormalities of the Microfibrillar-Fiber System in the Marfan Syndrome. N Engl J Med. 1990;323: 152–9. doi:10.1056/NEJM199007193230303

57. Hashimoto G, Inoki I, Fujii Y, Aoki T, Ikeda E, Okada Y. Matrix metalloproteinases cleave connective tissue growth factor and reactivate angiogenic activity of vascular endothelial growth factor 165. J Biol Chem. American Society for Biochemistry and Molecular Biology; 2002;277: 36288–95. doi:10.1074/jbc.M201674200

58. Whitelock JM, Murdoch AD, Iozzo R V, Underwood PA. The degradation of human endothelial cell-derived perlecan and release of bound basic fibroblast growth factor by stromelysin, collagenase, plasmin, and heparanases. J Biol Chem. 1996;271: 10079–86.

59. Mañes S, Llorente M, Lacalle RA, Gómez-Moutón C, Kremer L, Mira E, et al. The matrix metalloproteinase-9 regulates the insulin-like growth factor-triggered autocrine response in DU-145 carcinoma cells. J Biol Chem. 1999;274: 6935–45.

60. Tatti O, Vehvilainen P, Lehti K, Keskioja J. MT1-MMP releases latent TGF-β1 from endothelial cell extracellular matrix via proteolytic processing of LTBP-1. Exp Cell Res. 2008;314: 2501–2514. doi:10.1016/j.yexcr.2008.05.018

61. Hyytiäinen M, Penttinen C, Keski-Oja J. Latent TGF-β Binding Proteins: Extracellular Matrix Association and Roles in TGF-β Activation. Crit Rev Clin Lab Sci. Taylor & Francis; 2004;41: 233–264. doi:10.1080/10408360490460933

62. Bennike TB, Carlsen TG, Ellingsen T, Bonderup OK, Glerup H, Bøgsted M, et al. Neutrophil Extracellular Traps in Ulcerative Colitis: A Proteome Analysis of Intestinal Biopsies. Inflamm Bowel Dis. Wolters Kluwer Health; 2015;21: 2052–67. doi:10.1097/MIB.0000000000000460

63. Udenfriend S. Formation of hydroxyproline in collagen. Science. American Association for the Advancement of Science; 1966;152: 1335–1340. doi:10.1126/SCIENCE.152.3727.1335

64. Kessler S, Rho H, West G, Fiocchi C, Drazba J, de la Motte C. Hyaluronan (HA) deposition precedes and promotes leukocyte recruitment in intestinal inflammation. Clin Transl Sci. 2008;1: 57–61. doi:10.1111/j.1752-8062.2008.00025.x

65. Kugathasan S, Denson LA, Walters TD, Kim M-O, Marigorta UM, Schirmer M, et al. Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. Lancet. 2017;389: 1710–1718. doi:10.1016/S0140-6736(17)30317-3

66. Lu H, Hoshiba T, Kawazoe N, Chen G. Comparison of decellularization techniques for preparation of extracellular matrix scaffolds derived from three-dimensional cell culture. J Biomed Mater Res A. 2012;100: 2507–16. doi:10.1002/jbm.a.34150

67. Shimshoni E, Sagi I. Sample Preparation of Extracellular Matrix of Murine Colons for Scanning Electron Microscopy. In: Sagi I, Afratis N, editors. Collagen Methods in Molecular Biology. Humana Press, New York, NY; 2019. pp. 129–133. doi:10.1007/978-1-4939-9095-5_9

68. Butt H-J, Cappella B, Kappl M. Force measurements with the atomic force microscope: Technique, interpretation and applications. Surf Sci Rep. 2005;59: 1–152. doi:10.1016/j.surfrep.2005.08.003

69. Schillers H, Rianna C, Schäpe J, Luque T, Doschke H, Wälte M, et al. Standardized Nanomechanical Atomic Force Microscopy Procedure (SNAP) for Measuring Soft and Biological Samples. Sci Rep. 2017;7: 5117. doi:10.1038/s41598-017-05383-0

70. Cramér H. Mathematical Methods of Statistics. In: Princeton University Press. 1999.

71. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008;26: 1367–1372. doi:10.1038/nbt.1511

72. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry – based shotgun proteomics. Nat Protoc. Nature Publishing Group; 2016;11: 2301–2319. doi:10.1038/nprot.2016.136

73. Cox J, Neuhauser N, Michalski A, Scheltema R a., Olsen J V., Mann M. Andromeda: A peptide search engine integrated into the MaxQuant environment. J Proteome Res. 2011;10: 1794–1805. doi:10.1021/pr101065j

74. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Mol Cell Proteomics. 2014;13: 2513–2526. doi:10.1074/mcp.M113.031591

75. Salzberg SL. C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. Mach Learn. 1994;16: 235–240. doi:10.1007/BF00993309

76. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(- Delta Delta Ct) Method. Methods. 2001;25: 402–408. doi:10.1006/meth.2001.1262

77. Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers : Do not use standard deviation around the mean , use absolute deviation around the median. J Exp Soc Psychol. 2013; 4–6.

78. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods. Nature Research; 2016;13: 731–740. doi:10.1038/nmeth.3901

79. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci. 2001;98: 5116–5121. doi:10.1073/pnas.091062498

80. Solomonov I, Zehorai E, Talmi-Frank D, Wolf SG, Shainskaya A, Zhuravlev A, et al. Distinct biological events generated by ECM proteolysis by two homologous

collagenases. Proc Natl Acad Sci U S A. National Academy of Sciences; 2016;113: 10884–9. doi:10.1073/pnas.1519676113

1015

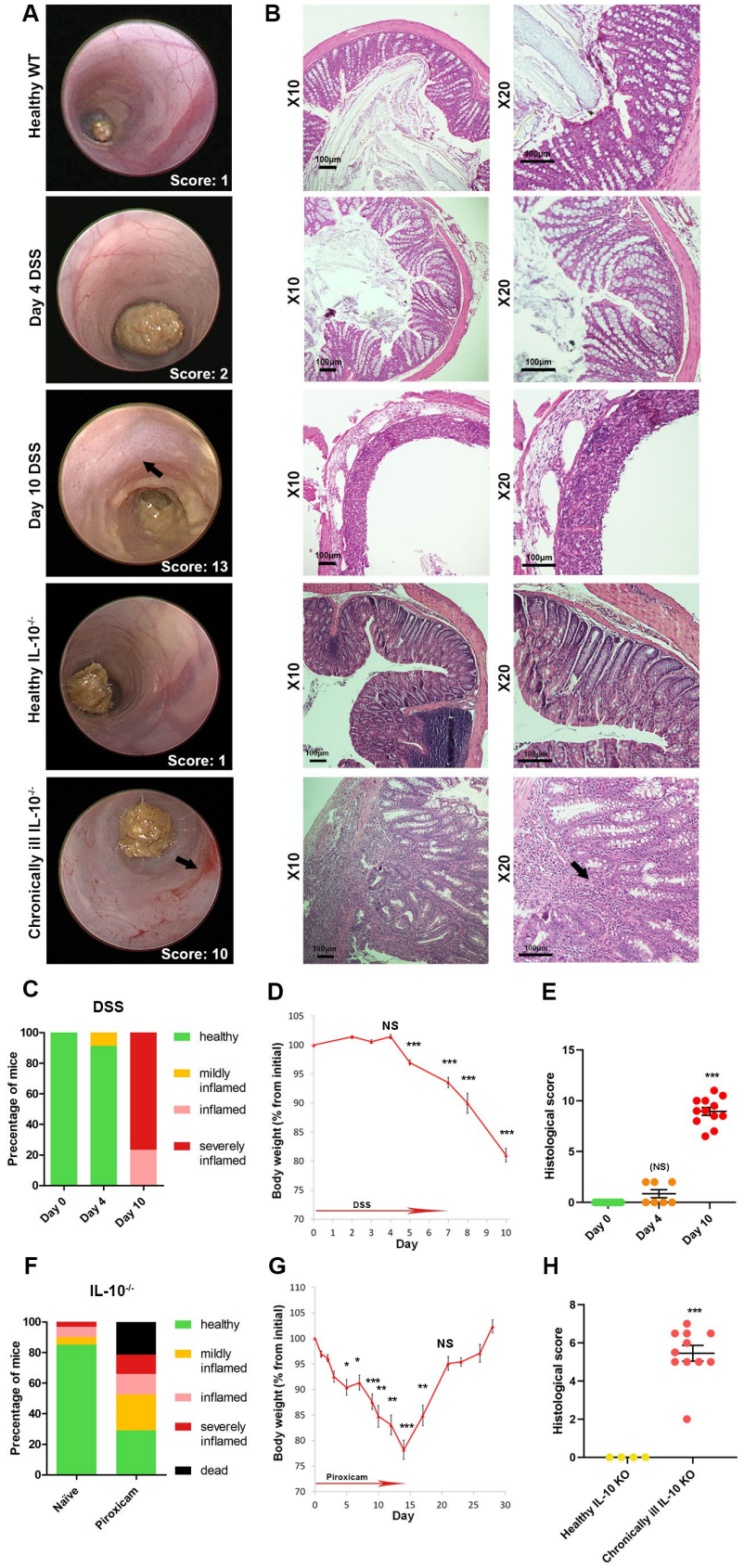
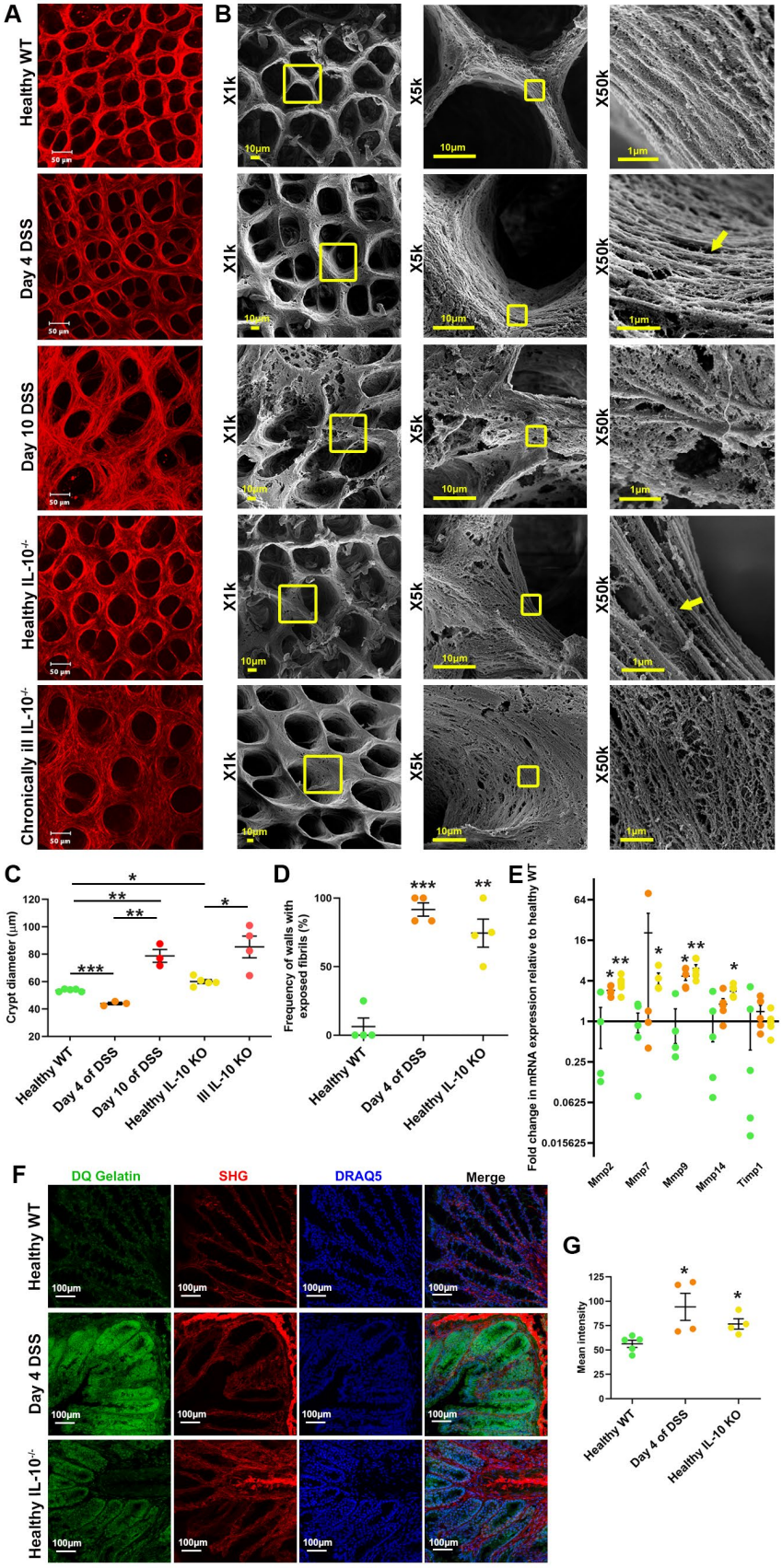

A
Healthy WT
Score: 1
Day 4 DSS
Score: 2
Day 10 DSS
Score: 13
Healthy IL-10-/-
Score: 1
Chronically ill IL-10-/-
Score: 10
B
X10
X20
100µm
C
DSS
Precentage of mice
healthy
mildly inflamed
inflamed
severely inflamed
Day 0
Day 4
Day 10
D
Body weight (% from initial)
NS
DSS
Day
E
Histological score
(NS)
Day 0
Day 4
Day 10
F
IL-10-/-
Precentage of mice
healthy
mildly inflamed
inflamed
severely inflamed
dead
Naïve
Piroxicam
G
Body weight (% from initial)
NS
Piroxicam
Day
H
Histological score
Healthy IL-10 KO
Chronically ill IL-10 KO

Fig.1

Acute model – DSS-induced colitis. Clinical and endoscopic symptoms are apparent from day 7 and peak at day 10. Chronic model – PAC IL-10$^{-/-}$ model. Clinical and endoscopic symptoms develop over the course of these 14 days, and the chronic inflammatory state is persistent per mouse. Chronically ill IL-10$^{-/-}$ mice were harvested at least two weeks following piroxicam discontinuation. (A) Endoscopic images. Live mice were imaged at the indicated states of the two models and inflammation scored as described in Materials and Methods section (0-4: healthy; 5-7: mildly inflamed; 8-11: inflamed; 12-15: severely inflame). The arrow in image of day 10 of the DSS model indicates the thick, granulated opaque mucosal surface. Chronically ill IL-10$^{-/-}$ mice predominantly display vascular disturbances and bleeding, as indicated by arrow. (B) H&E-stained colonic sections of mice from the two models at the indicated states. Note that immune cell infiltration and mucosal damage is not evident on day 4 of the acute model, but is substantial on day 10 (indicated by arrow). In the PAC IL-10$^{-/-}$ model large amounts of immune cell infiltrate is observed in the mucosa (indicated by arrow), but is absent in the healthy IL-10$^{-/-}$ mouse. (C+F) Quantification of the percentage of mice in each clinical category at the indicated time points or states of both models (n>30). Note that some IL-10$^{-/-}$ mice develop spontaneous inflammation without piroxicam exposure, and that some do not develop inflammatory symptoms following exposure. (D+G) Mouse body weight changes over the course of the two models, presented as percentage from body weight on day 0. Significant weight loss in the DSS model appears from day 5 compared to day 2. Number of animals: n(2)= 85, n(4)=55, n(5)=36, n(7)=41, n(8)=10, n(10)=52; Bonferroni for five comparisons: ***P<0.0002. Significant weight loss appears in the IL-10$^{-/-}$ from day 5 compared to day 1. Number of animals: n(1)= 8, n(5)=22, n(7)=7, n(9)=7, n(10)=13, n(12)=19, n(14)=10, n(17)=10, n(21)=17; Bonferroni correction for

1040 eight comparisons: *P<0.00625, **P<0.00125, ***P<$1.25X10^{-4}$. Note that mice regain their weight as they reach the chronic phase of colitis. (E+H) Histological scoring of H&E stained sections, such as the ones shown in B. Details on scoring are available in the Materials and Methods section. In the DSS model, a significant increase in histopathological score was only observed on Day 10 and not on Day 4 (NS=not significant). Chronically ill IL-$10^{-/-}$ display a

1045 significant increase in histopathological score. The plot displays the scores of individual animals and the bars represent the state's mean±SEM. Statistical significance was determined using a student's t-test with a Bonferroni correction for two comparisons to healthy WT in E (NS: P> 0.025, *** P<0.0005) and with no correction, compared to healthy IL-$10^{-/-}$ in H (*** P<0.001).


A
B
Healthy WT
Day 4 DSS
Day 10 DSS
Healthy IL-10-/-
Chronically ill IL-10-/-
X1k
X5k
X50k
C
Crypt diameter (µm)
Healthy WT
Day 4 of DSS
Day 10 of DSS
Healthy IL-10 KO
Ill IL-10 KO
D
Frequency of walls with exposed fibrils (%)
E
Fold change in mRNA expression relative to healthy WT
Mmp2
Mmp7
Mmp9
Mmp14
Timp1
F
DQ Gelatin
SHG
DRAQ5
Merge
G
Mean intensity

1050 **Fig. 2 The ECM suffers structural damage involving gelatinase activity during colitis**

(A) Second-harmonic imaging of mouse colon on the indicated time points over the course of the disease, corresponding to distinct states: healthy WT, pre-symptomatic (day 4 of DSS and healthy IL-10$^{-/-}$), acute inflammation (day 10 of DSS) and chronic inflammation (Chronically ill IL-10$^{-/-}$). Vast ECM structural changes occur in acute inflammation, as indicated in the

1055 comparison between day 0 and day 10. Remarkably, some reorganization of the ECM is apparent from day 4, before clinical symptoms are evident. Healthy IL-10$^{-/-}$ display different ECM structure than that of WT mice, with ECM condensation similar to that of day 4. Chronic illness leads to overall maintained crypt architecture, but with loosely packed ECM. (B) SEM images of decellularized colonic ECM at the corresponding states, at three different magnifications. Note

1060 the damage and crypt size heterogeneity apparent on day 10 compared to Healthy WT colon. Also note, that the ECM under chronic illness is characterized by perforated crypt walls with thin fibrils. Most remarkably, the two pre-symptomatic states show a common feature of exposure of fibrillar ECM proteins on the crypt walls, as can be observed in the X50k magnification (examples are indicated by arrows). (C) Measurement of crypt diameter in second-harmonic

1065 images. In oval crypts, the largest diameter was always chosen. Note each state has a different mean diameter, indicating architectural changes of the ECM. Statistical significance was determined using a t-test with a Bonferroni correction for five comparisons, and is indicated by asterisks. *P<0.01, **P<0.002, ***P<0.0002. The bars represent the mean±SEM of the state, while each dot corresponds to the mean diameter of one animal, which is based on the analysis of

1070 2-8 440μmX440μm images capturing n>30 crypts per animal. One outlier was removed in the Day 4 group. (D) The frequency of exposed fibrils on crypt walls in the SEM images, indicating a leaky ECM underlying the epithelium, in Healthy WT vs. Day 4 of DSS and Healthy IL-10$^{-/-}$

samples. **P<0.005, ***P<0.0005, n=4 different samples per state, based on analysis of n>14 crypts. (E) Gene expression analysis of selected ECM remodeling enzymes (Mmp-2/-7/-9/-14) and inhibitor (TIMP1) at pre-symptomatic states. An independent set of mouse tissue samples of Healthy WT (green), Day 4 of DSS (orange) and Healthy IL-10$^{-/-}$ (yellow) were used. Note, that gelatinases (Mmp-2/-9) are significantly upregulated at pre-symptomatic states. n=5 animals per state; *P<0.025, **P<0.005, t-test with Bonferroni correction. (F) In situ zymography with DQ Gelatin™ on colon cross-sections, depicting elevated gelatinase activity (green) in the mucosa of the pre-symptomatic states, along with DRAQ5-stained cell nuclei (blue) and SHG signal of fibrillary collagen (red). (G) Quantification of signal intensity in the mucosa, revealing that the elevation in gelatinase activity is statistically significant.

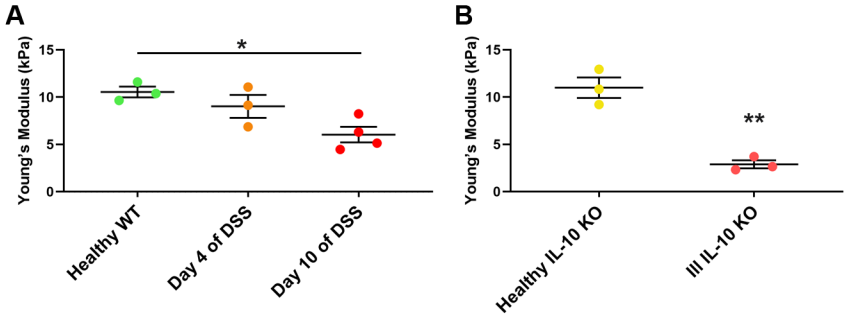


Fig. 3 The ECM becomes softer during colitis

(A+B) ECM softening during colitis progression: individual dots represent the properly weighted median values of the Young's modulus for each animal and the bars represent the average median±STD, corresponding to all distributions represented in **Fig. S2**. (A) The results show a significant reduction in ECM rigidity between the healthy and acutely inflamed samples as denoted by the asterisks (according to a t-test with Bonferroni correction for three comparisons, *P<0.017). The differences between the pairs HC-D4 and D4-D10 are not statistically significant (P=0.30 and P=0.09, respectively); nevertheless, pre-symptomatic ECMs (D4) show an intermediate behavior with respect to the healthy and acutely inflamed cases. (B) The results show a significant reduction in ECM rigidity between the healthy and chronically inflamed states, as denoted by the asterisks (**P<0.01).

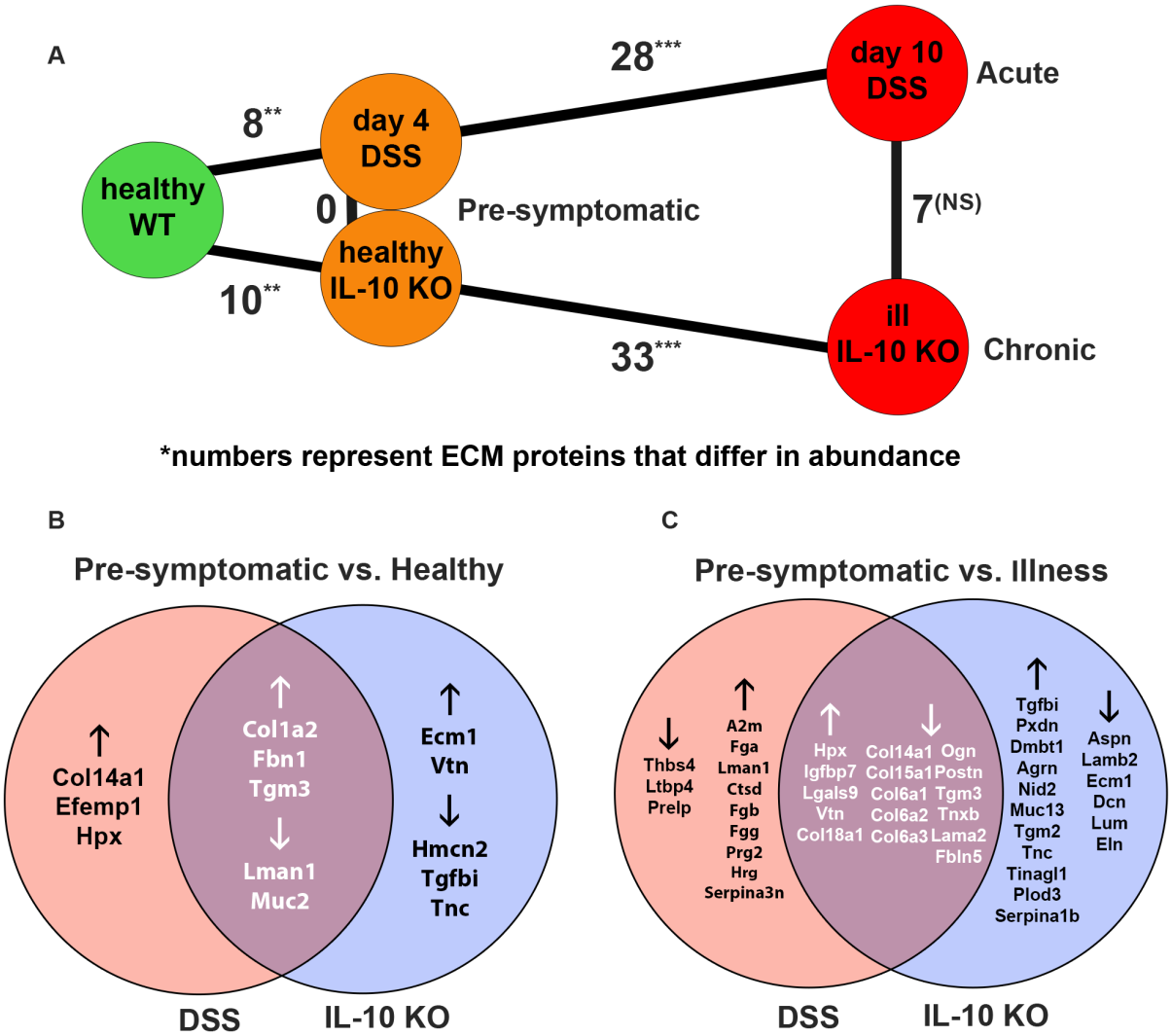


Fig. 4 ECM compositional differences and similarities during colitis are in concordance with structural changes, as revealed by mass-spectrometry (MS) analysis

(A) Scheme showing the number of differentially abundant ECM proteins, between pairs in the five different states, which is represented by lines stretched between two nodes of two states. Note that the two pre-symptomatic states are similar in terms of ECM composition. Statistical significance of amount of ECM proteins changing among all ECM proteins and all differentially abundant proteins, according to the sum of hypergeometric probabilities for achieving each

number or above it, is indicated next to each comparison by asterisks. **P<0.01, *** P<0.001, NS=not significant. (B) Venn diagram presenting the differentially abundant ECM proteins in the colon comparing healthy WT vs. the pre-symptomatic day 4 of the DSS model (pink) and healthy WT vs. pre-symptomatic healthy IL-10$^{-/-}$ (blue). Direction of change in abundance in pre-symptomatic states is indicated by arrows. Note the five differentially abundant proteins shared among the two comparisons (purple). Intersection between the differentially abundant proteins in both comparisons is statistically significant ($p=1.0x10^{-4}$). (C) Venn diagram presenting the amounts of differentially abundant ECM proteins in the colon comparing pre-symptomatic day 4 of the DSS model vs. acutely ill day 10 of the DSS model (pink) and pre-symptomatic healthy IL-10$^{-/-}$ vs. chronically ill IL-10$^{-/-}$ (blue). Note the large portion of differentially abundant proteins shared among the two comparisons (purple). Intersection between the differentially abundant proteins in both comparisons is statistically significant ($p=1.2x10^{-5}$). Statistical significance in this figure was calculated by the sum of hypergeometric probabilities for achieving each number or above it.

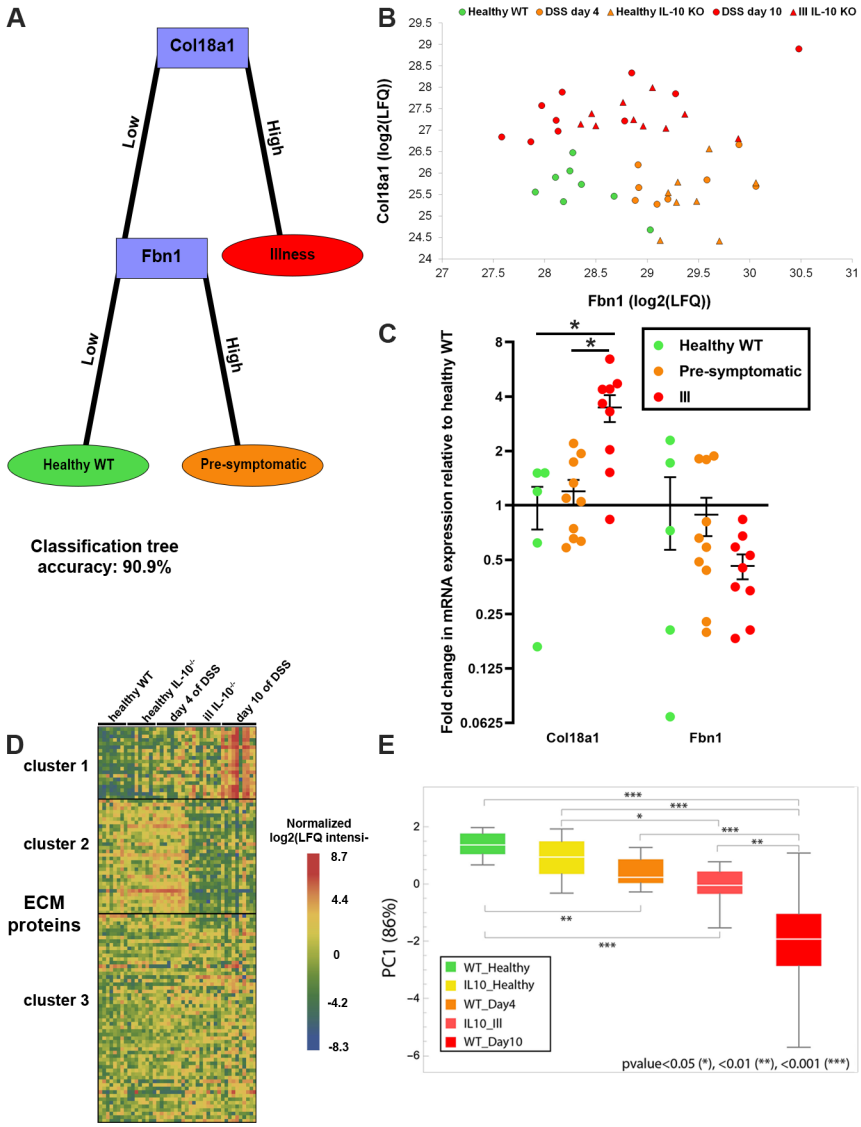


Fig. 5 Machine-learning-based analysis of ECM composition decodes composition and unravels potentially pivotal components

(A) Method for identifying potential biomarkers by applying a classification algorithm. Samples were divided into Healthy WT, Pre-symptomatic (day 4 of DSS and healthy IL-10$^{-/-}$) and Ill (day 10 DSS and ill IL-10$^{-/-}$), and a J48 decision-tree classification algorithm[75] was applied to the data, along with a stratified 10-fold cross-validation on Weka software. The two proteins chosen for the tree nodes (Col18a1 and Fbn1) are those that their abundance separate the three groups with the highest accuracy (90.9%) of all proteins. Col18a1 threshold: 26.6637 log2(LFQ

1130 intensity). Fbn1 threshold: 28.76329 log2(LFQ intensity). (B) Plot of all samples according only to the abundances of Col18a1 and Fbn1, demonstrating that the abundances of the two proteins are good indicators of the tissue state and good separators of the data. (C) Gene expression analysis of Collagen XVIII (Col18a1) and Fibrillin-1 (Fbn1) on independent mouse tissue samples of all three groups. Note, that Col18a1 transcription is significantly elevated in ill

1135 samples, while no significant changes were observed in Fbn1. n(Healthy WT)=5, n(Pre-symptomatic)=10, n(Ill)=9 (one outlier removed); *P<0.0167, t-test with Bonferroni correction. (D) MS data of 110(ECM proteins) x 44(samples at five clinical states) was clustered according to similar protein abundance by computing the squared Euclidean distance. Following, the data was partitioned into three clusters. Grouping according to

1140 clinical state is indicated on the x axis. Protein abundances were normalized to the mean of each protein. (E) PCA on the 3(clusters) x 44(samples) data after taking the mean of each cluster. The first PC explains more than 86% of the variance in the data. PC1 mostly (~75%) consists of cluster 1. The PC1 values for each state are presented in a box plot ranging from the first and third quartiles, with the median indicated by a white line. Whiskers represent the minimum and

1145 maximum values. *P<0.05, **P<0.01, *** P<0.001. Note that the results show that ECM composition of clinical states can be projected onto a healthy-acutely ill axis.

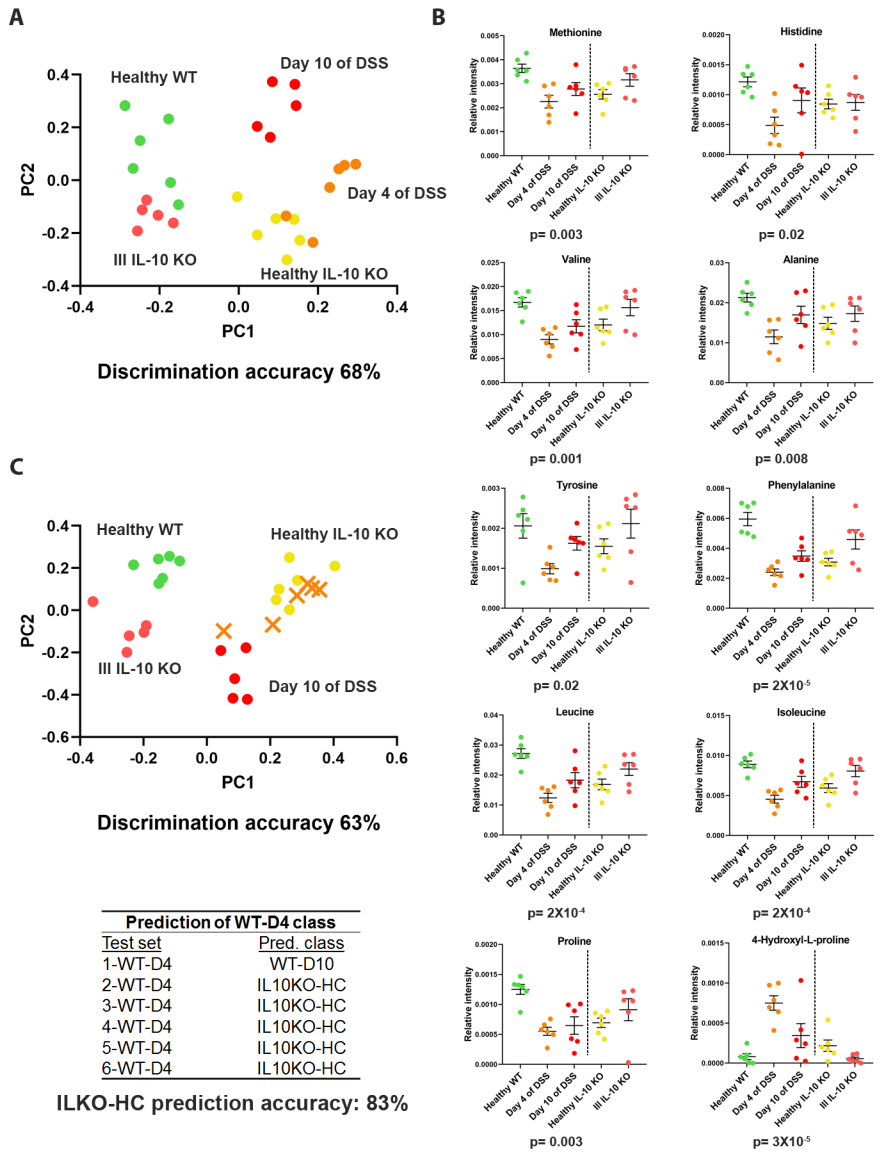


Fig. 6 Extracellular matrix proteolysis reveals its role as a metabolite reservoir that is modified under different inflammatory conditions

1150 NMR-based untargeted metabolomic analysis on solutions with products of MMP-1 proteolysis of decellularized murine colon tissues in the five different states – WT-HC (green), WT-D4 (orange) IL10-HC (yellow) dots, WT-D10 (red), IL10-ILL (pink). (A) PCA-CA analysis revealed that MMP-1 proteolysis resulted in differentiating metabolic profiles of the degradation products across the five colitis states of the two models, with a 5-group discrimination accuracy

1155 of 68%. (B) Scatter plots of metabolites displaying statistically significant different levels in the five states with the indicated one-way ANOVA p values. Pre-symptomatic states (WT-D4 and IL10-HC) display reduced extraction of amino acids upon proteolysis compared to both WT-HC and the two ill states (WT-D10 and IL10-ILL). This indicates that the pre-symptomatic ECM is a poor reservoir of these types of metabolites. However, most interestingly, WT-D4 matrix
1160 releases high amounts of 4-Hydroxyl-L-proline, the unique amino acid found in collagen fibrils, indicating the susceptibility of the collagenous ECM at this pre-symptomatic state to degradation. (C) PCA-CA prediction of WT-D4 group (orange crosses) using a training set composed of the other four groups. WT-D4 samples are predicted as the other pre-symptomatic state – IL10-HC). In the score plots each symbol represents a different sample and each color a
1165 different group of mice.

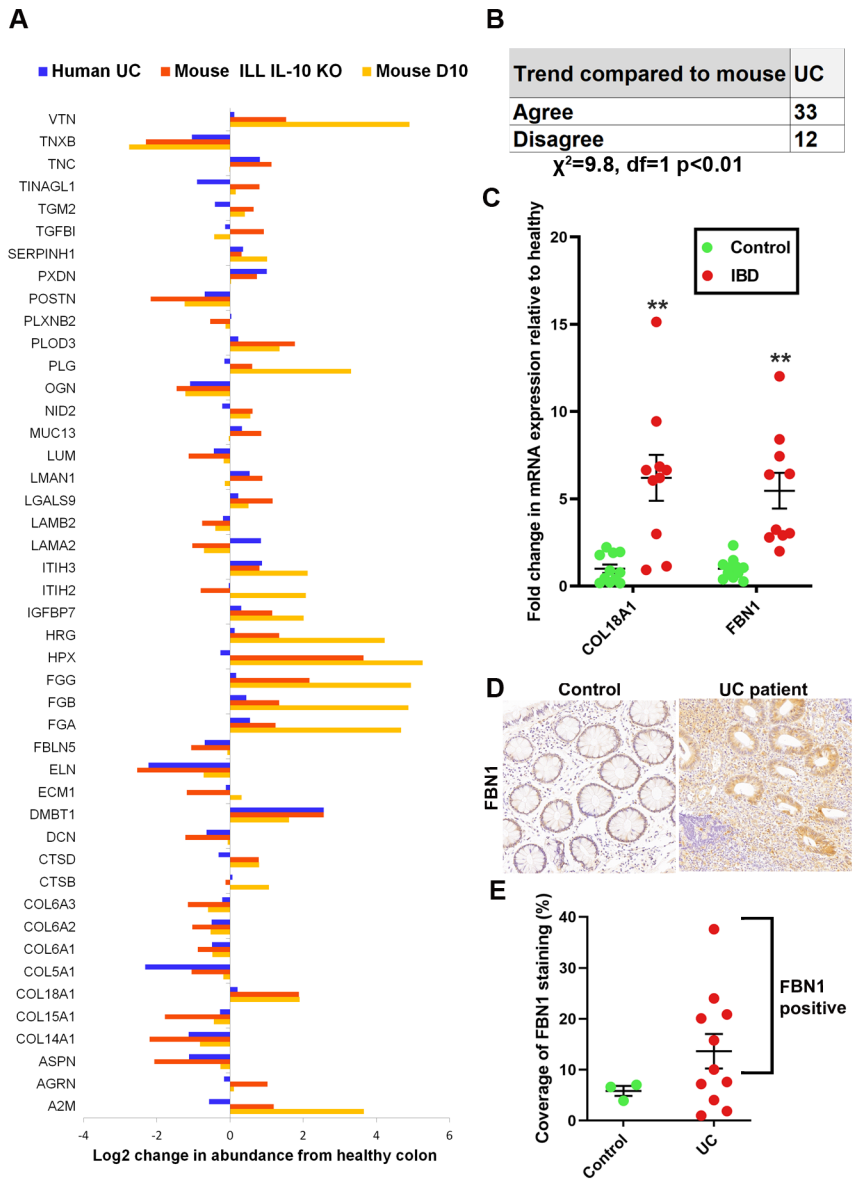


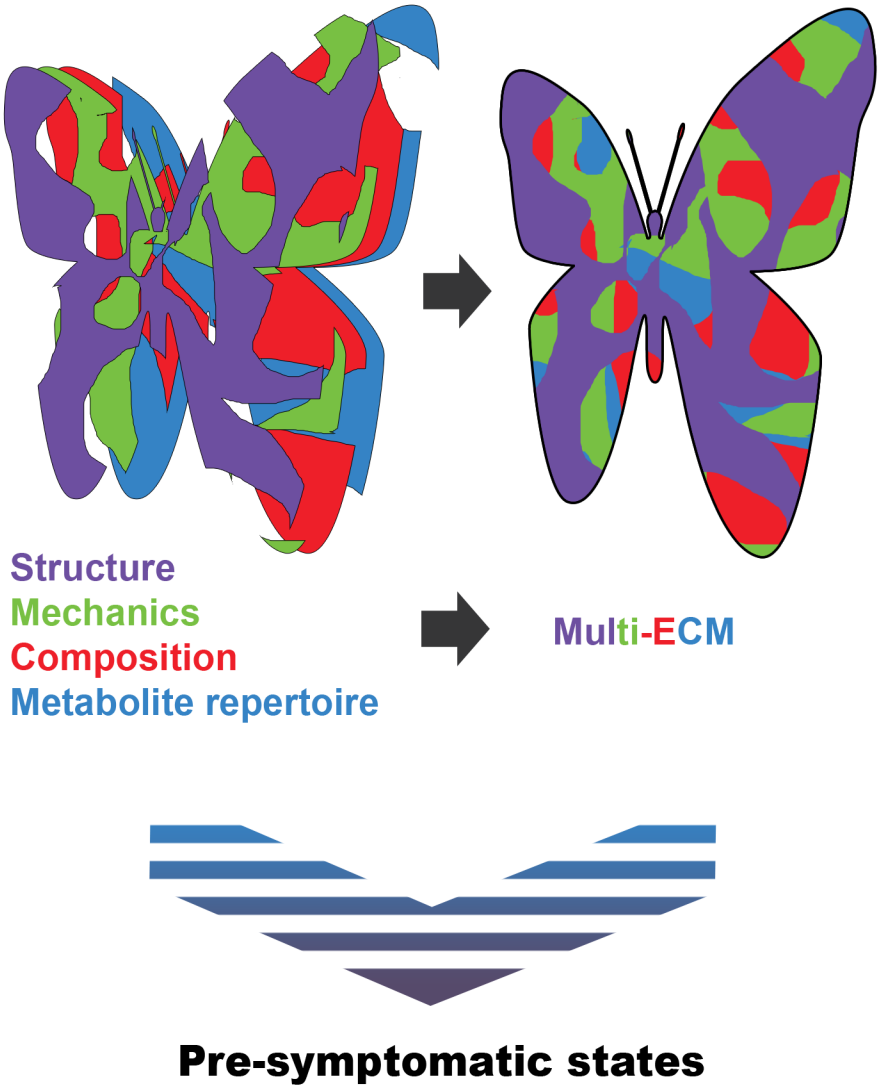
Fig. 7 Human IBD patients display similar ECM proteomic signatures

(A) Significantly different ECM proteins from MS analysis display similar trends in human

1170 dataset. Unpublished data from a proteomic data set of colon biopsies, Bennike *et al.*[62] were analyzed for changes in ECM protein abundance, comparing UC[62] patients to healthy controls. Proteins that changed significantly between healthy (WT or IL-10$^{-/-}$) and ill (day 10 of DSS or chronically ill IL-10$^{-/-}$, respectively) mice were chosen. (B) The directions of the observed changes in human were compared to that of the significant changes in mouse. Proteins that

1175 change in the same direction in both species were counted as those that “Agree” with the mouse data, compared to those that “Disagree”. A chi-square test of goodness-of-fit was performed to determine whether the rate of agreement in the direction of change was likely to occur by chance by comparing to a 50% (22.5) “Agree”-50% (22.5) “Disagree” distribution. (C) Validation of biomarker transcript levels on an independent set of human samples. Biopsies and surgical
1180 resections of healthy donors or of inflamed mucosa of IBD patients were analyzed for mRNA expression levels of COL18A1 and FBN1. Both predicted markers indicated by the classification model performed on the murine models display significantly elevated levels of 6.2 fold (COL18A1) and 5.5 fold (FBN1) in inflamed IBD tissue compared to tissue from healthy donors (n(IBD)=10, n(Healthy)=11; *P<0.05, **P<0.01). (D) FBN1 IHC staining in rectal biopsies of
1185 healthy donors vs. inflamed regions in UC patients. Controls do not demonstrate FBN1 staining (FBN1 negative), while a subset of UC patients display a strong epithelial staining (FBN1 positive). (E) Quantification of the area covered by FBN1 staining. The distribution of the FBN1 coverage in UC patients is broad, and half of the patients analyzed are FBN1 positive (above 10% coverage), such as the one presented in D.

1190



Scheme 1. Applying multidimensional ECM analysis provides a complete picture of ECM dynamics throughout pathogenesis and leads to novel findings.

Supplementary Information for

Multidimensional extracellular-matrix analysis reveals a silent pre-clinical signature in inflammatory disease

Elee Shimshoni[1], Ran Afik[1], Anjana Shenoy[2], Miri Adler[3], Luca Puricelli[4], Veronica Ghini[5], Odelia Mouhadeb[6,8], Nathan Gluck[6,8], Sigal Fishman[6,8], Lael Werner[7,8], Dror S. Shouval[7,8], Chen Varol[6,8], Alessandro Podestà[4], Paola Turano[5], Tamar Geiger[2], Paolo Milani[4], Claudio Luchinat[2], Uri Alon[3] and Irit Sagi[1*].

Correspondence to: irit.sagi@weizmann.ac.il

This includes:

Figures S1 to S3
Tables S1 to S7

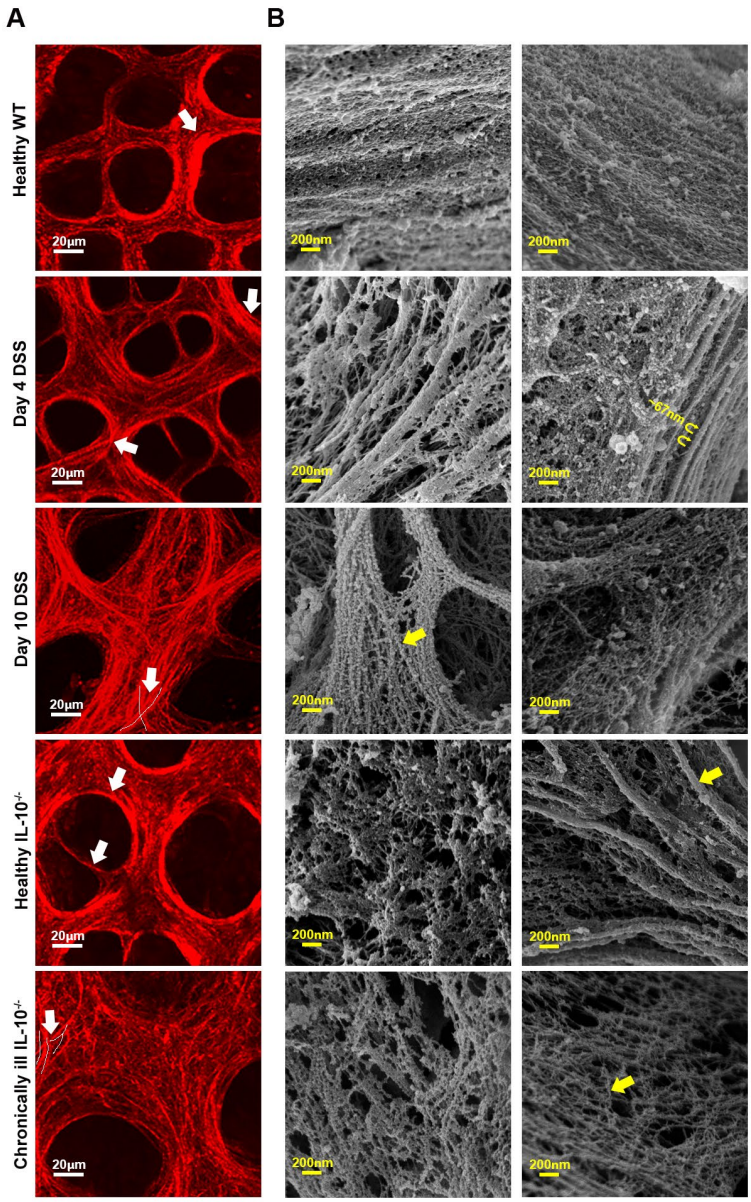
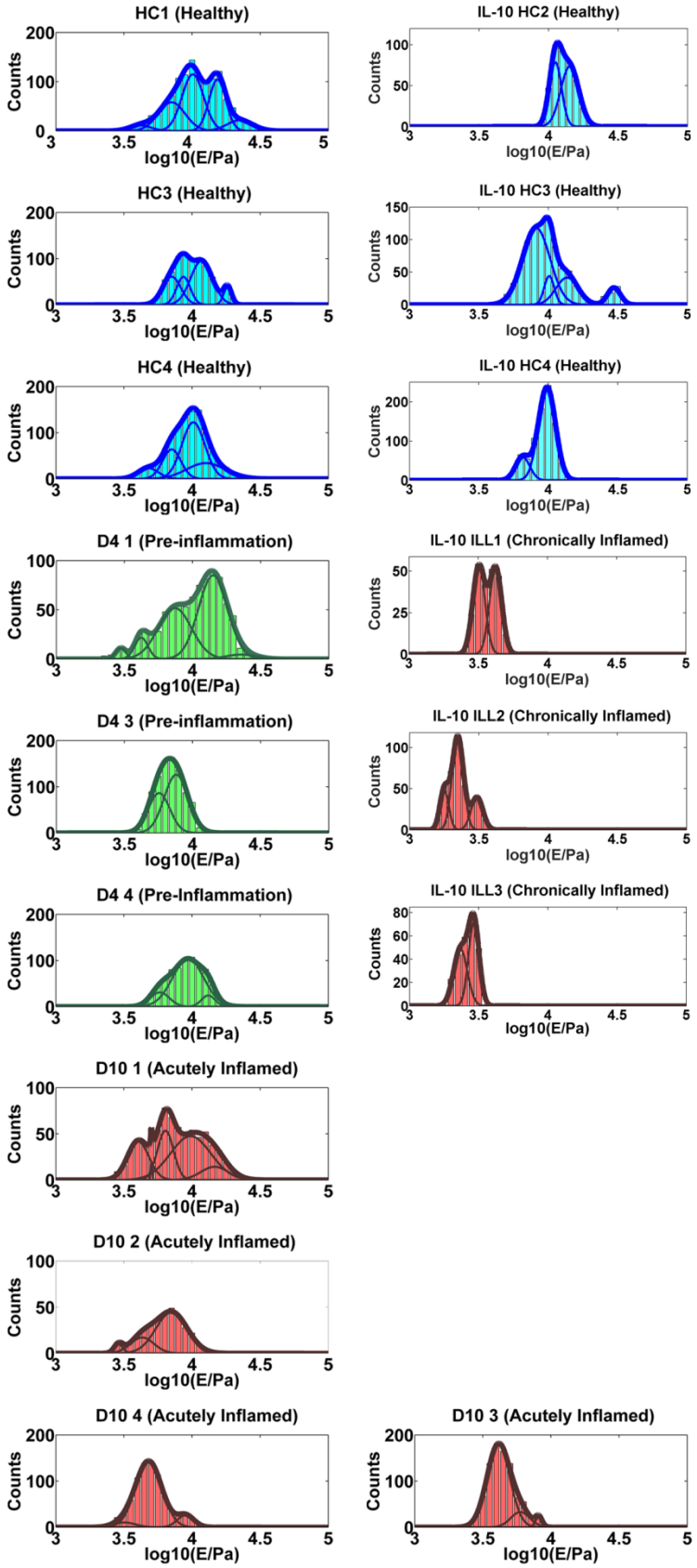


Fig. S1.

High-magnification images of SHG and SEM for emphasizing the structural changes and features of the ECM in all five tissue states. (A) Zoomed-in images of the ones in **Fig. 2A**. The higher magnification images allow us to demonstrate the changes in collagen fibers, which are pointed out by arrows. In healthy WT mice collagen is condensed around crypt borders, while fibers are somewhat splitting on day 4 of the DSS model as well as in healthy IL-10 mice. We

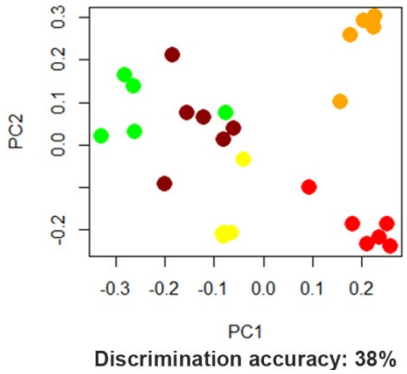
see how the fibers become even more disoriented in both ill states, which is highlighted by white lines that follow selected intersecting fibers. This indicates that collagen fibers are splitting as a result of enzymatic degradation [80]. The high magnification also emphasizes the loose collagen structure characteristic of the chronic illness that contains thinner fibrils, which do not emit a strong SHG signal. In contrast, day 10 of the DSS model has fibers that emit a strong signal, indicating the coincidence of degradation (splitting) and deposition. (B) Additional SEM images, each image came from a different mouse. The images demonstrate the deterioration observed when comparing healthy WT to the two pre-symptomatic states (day 4 of DSS and healthy IL-10) in the ECM network that comprises the first layer that underlies the epithelium, which is most likely the basement membrane, since it resembles a collagen type IV network[44]. Not only are collagen fibrils exposed, as indicated by their characteristic 67nm D-banding (pointed out on day 4 in this panel), but also the network that coats them is not condensed and contains holes, both at day 4 and in healthy IL-10 mice. In addition, the changes in the observed fibrils as illness develops are pointed out by the arrows – while we see thick fibers at pre-symptomatic states, those are not present in acute (day 10) or chronic colitis (ill IL-10), and only thin, disoriented fibrils can be observed.


HC1 (Healthy)
IL-10 HC2 (Healthy)
HC3 (Healthy)
IL-10 HC3 (Healthy)
HC4 (Healthy)
IL-10 HC4 (Healthy)
D4 1 (Pre-inflammation)
IL-10 ILL1 (Chronically Inflamed)
D4 3 (Pre-inflammation)
IL-10 ILL2 (Chronically Inflamed)
D4 4 (Pre-Inflammation)
IL-10 ILL3 (Chronically Inflamed)
D10 1 (Acutely Inflamed)
D10 2 (Acutely Inflamed)
D10 4 (Acutely Inflamed)
D10 3 (Acutely Inflamed)
Counts
log10(E/Pa)

1240

Fig. S2.

Collective distributions of Young's Modulus values (E) measured on ECMs derived from colons of mice at three different states of the DSS-induced colitis model and two states of the PAC IL-$10^{-/-}$ model (HC: healthy colon; D4: pre-symptomatic; D10: acutely inflamed; IL-10 HC: healthy naïve IL-$10^{-\backslash-}$; IL-10 ILL: chronically inflamed IL-$10^{-/-}$). Each sample was derived from a different mouse. Thick lines represent the result of the superposition of the underlying major contributions, highlighted by thin continuous lines. Single contributions are detected by fitting the envelope of several Gaussian profiles to data in semilog10 scale. Under the hypothesis that the underlying distributions are log-normal, the peaks of the Gaussians represent the median values of the Young's Modulus. Different contributions arise from the fact that each ECM is tested in different locations, therefore multiple modes witness the mechanical and structural heterogeneity of ECMs.



1255 **Fig. S3.**

Decellularized murine colon tissues of the five different colitis states – WT-HC (green), WT-D4 (yellow) IL10-HC (orange), WT-D10 (red), IL10-ILL (dark red), were incubated in TNC buffer for 24 hours at 37˚C. The solution was subjected to untargeted metabolomic analysis using NMR. (A) PCA-CA analysis revealed that the metabolic repertoires were not distinguishable

1260 without MMP addition (38% discrimination accuracy).

Sample		Young's Modulus / kPa (weighted median ± propagated errors)	Young's Modulus / kPa (average median ± effective STD of the mean)
Healthy WT (day 0)	HC 1	11.61 ± 0.25	10.55 ± 0.49
	HC 3	10.39 ± 0.26	
	HC 4	9.65 ± 0.25	
Pre-Symptomatic (day 4)	D4 1	11.07 ± 0.39	9.03 ± 1.01
	D4 3	6.86 ± 0.25	
	D4 4	9.15 ± 0.36	
Acutely inflamed (day 10)	D10 1	8.23 ± 0.24	6.04 ± 0.72
	D10 2	6.31 ± 0.24	
	D10 3	4.47 ± 0.15	
	D10 4	5.14 ± 0.18	
Healthy IL-10$^{-/-}$	IL-10 HC2	12.94 ± 0.30	11.00 ± 0.90
	IL-10 HC3	10.84 ± 0.20	
	IL-10 HC4	9.21 ± 0.24	
Chronically ill IL-10$^{-/-}$	IL-10 ILL1	3.71 ± 0.06	2.89 ± 0.34
	IL-10 ILL2	2.32 ± 0.03	
	IL-10 ILL3	2.64 ± 0.04	

Table S1.

The statistical indicators and associated errors derived from the single-mode analysis for the collective distributions of Young's Modulus values of ECMs represented in **Fig. S2**, which are plotted in **Fig.3A-B**. P values when comparing average medians: Healthy WT vs. Day 10 (P<0.01); the differences between Healthy WT vs. Day 4 and Day 4 vs. Day 10 are not statistically significant (P=0.30 and P=0.09, respectively); Healthy IL-10$^{-/-}$ vs. Ill IL-10$^{-/-}$ (P<0.002). STD – standard deviation. All data are expressed in kPa units.

Sample	Young's Modulus / kPa (median ± STD of the median)					
Healthy WT 1		4.17 ± 0.14	7.08 ± 0.12	10.12 ± 0.11	15.33 ± 0.15	22.40 ± 0.60
Healthy WT 3			7.08 ± 0.19	8.63 ± 0.10	18.20 ± 0.21	
Healthy WT 4		4.76 ± 0.11	7.08 ± 0.09	10.20 ± 0.11 12.60 ± 0.33		
Day 4 1	2.98 ± 0.07	4.21 ± 0.08	7.42 ± 0.15		14.16 ± 0.20	22.66 ± 1.46
Day 4 3			5.68 ± 0.07 7.59 ± 0.08			
Day 4 4			5.80 ± 0.11	9.35 ± 0.13	13.16 ± 0.24	
Day 10 1		4.02 ± 0.07	5.03 ± 0.03 6.38 ± 0.09	9.76 ± 0.22	14.64 ± 0.50	
Day 10 2	2.89 ± 0.07	4.31 ± 0.12	7.08 ± 0.15			
Day 10 3		4.11 ± 0.04	5.96 ± 0.11	8.12 ± 0.11		
Day 10 4	3.19 ± 0.11	4.81 ± 0.05			8.96 ± 0.18	
IL-10 Healthy 2				11.19 ± 0.11 14.20 ± 0.17		
IL-10 Healthy 3			8.17 ±0.10	10.13 ± 0.12 13.62 ± 0.24	29.66 ± 0.51	
IL-10 Healthy 4			6.52 ± 0.08	9.77 ± 0.06		
IL-10 Ill 1		3.20 ± 0.04 4.23 ± 0.05				
IL-10 Ill 2	1.78 ± 0.02 2.23 ± 0.02	3.07 ± 0.04				
IL-10 Ill 3	2.33 ± 0.03	2.91 ± 0.03				

Table S2.

Major contributions to the broad statistical distributions of the Young's modulus values (see Materials and Methods section for details). STD – standard deviation. All data are expressed in kPa units.

Comparison:	HCvsD4	HCvsIL10H	D4vsIL10H	D4vsD10	IL10HvsIL10I	D10vsIL10I
ECM/Total	8/52	10/79	0/16	28/290	33/296	7/81
List of ECM proteins:	Col14a1	Col1a2		A2m;Pzp	Agrn	Anxa5
	Col1a2	Ecm1		Col14a1	Aspn	Aspn
	Efemp1	Fbn1		Col15a1	Col14a1	Serpina1a
	Fbn1	Hmcn2		Col18a1	Col15a1	Serpina1e
	Hpx	Lman1		Col6a1	Col18a1	Serpina3k
	Lman1	Muc2		Col6a2	Col6a1	Serpina3n
	Muc2	Tgfbi		Col6a3	Col6a2	Tinagl1
	Tgm3	Tgm3		Ctsd	Col6a3	
		Tnc		Fbln5	Dcn	
		Vtn		Fga	Dmbt1	
				Fgb	Ecm1	
				Fgg	Eln	
				Hpx	Fbln5	
				Hrg	Hpx	
				Igfbp7	Igfbp7	
				Lama2	Lama2	
				Lgals9	Lamb2	
				Lman1	Lgals9	
				Ltbp4	Lum	
				Ogn	Muc13	
				Postn	Nid2	
				Prelp	Ogn	
				Prg2	Plod3	
				Serpina3n	Postn	
				Tgm3	Pxdn	
				Thbs4	Serpina1b	
				Tnxb	Tgfbi	
				Vtn	Tgm2	
					Tgm3	
					Tinagl1	
					Tnc	
					Tnxb	
					Vtn	

Down on right state

Up on right state

Table S3.

List of differentially abundant ECM proteins arising from MS results in comparisons of different state pairs. All differentially abundant proteins were identified by applying a t-test with a p-value cut off of 0.05 and permutation-based False Discovery Rate (FDR) correction. The second row indicates the proportion of ECM proteins out of the total number of differentially abundant proteins. Red indicates proteins that are more abundant on the state written on the right, and blue indicates proteins that are less abundant on the state written on the right. Abbreviations: HC=Healthy WT; D4=Day 4 of DSS model; D10=Day 4 of DSS model; IL10H=Healthy IL-10 mice; IL10I=Chronically ill IL-10 mice.

Cluster 1	Name	Details
A2m;Pzp	Alpha-2-Macroglobulin	Protease inhibitor
Anxa1	Annexin A1	Annexin
Fga	Fibrinogen alpha chain	Fibrinogen complex; coagulation
Fgb	Fibrinogen beta chain	Fibrinogen complex; coagulation
Fgg	Fibrinogen gamma chain	Fibrinogen complex; coagulation
Hpx	Hemopexin	Heme binding
Hrg	Histidine-rich glycoprotein	Coagulation
Itih2	Inter-alpha-trypsin inhibitor heavy chain H2	Serine-protease inhibitor
Itih3	Inter-alpha-trypsin inhibitor heavy chain H3	Serine-protease inhibitor
Itih4	Inter-alpha-trypsin inhibitor heavy chain H4	Serine-protease inhibitor
Kng1	Kininogen 1	Thiol-protease inhibitor; coagulation
Plg	Plasminogen	Protease
Serpina1a	Serine protease inhibitor A1A	Serine-protease inhibitor
Serpina1b	Serine protease inhibitor A1B	Serine-protease inhibitor
Serpina1d	Serine protease inhibitor A1D	Serine-protease inhibitor
Serpina1e	Serine protease inhibitor A1E	Serine-protease inhibitor
Serpina3k	Serine protease inhibitor A3K	Serine-protease inhibitor
Serpina3n	Serine protease inhibitor A3N	Serine-protease inhibitor
Serpinc1	Serine protease inhibitor C1	Serine-protease inhibitor
Vtn	Vitronectin	ECM glycoprotein
Cluster 2	Name	Details
Hmcn2	Hemicentin 2	ECM glycoprotein
Col15a1	Collagen, type XV, alpha 1 chain	Basement membrane
Col6a4	Collagen, type VI, alpha 4 chain	Microfibrillar collagen
Lgals4	Galectin-4	β-galactoside sugar binding
Aspn	Asporin	Small leucine-rich proteoglycan
Col5a1	Collagen, type V, alpha 1 chain	Fibrillary collagen
Thbs4	Thrombospondin 4	Antiangiogenic
Col14a1	Collagen, type XIV, alpha 1 chain	Fibril associated collagen with interrupted triple helices
Ltbp4	Latent TGFβ Binding Protein 4	Binds latent TGFβ
Col6a3	Collagen, type VI, alpha 3 chain	Microfibrillar collagen
Col12a1	Collagen, type XII, alpha 1 chain	Fibril associated collagen with interrupted triple helices
Vcan	Versican	ECM proteoglycan
Clec2d	C-type lectin domain family 2 member D	Cell surface receptor
Lamc1	Laminin subunit gamma 1	Basement membrane glycoprotein
Serpinb6	Serine protease inhibitor B6	Serine-protease inhibitor
Tnxb	Tenascin XB	Tenascin glycoprotein member; anti-adhesive
Col1a1	Collagen, type I, alpha 1 chain	Fibrillary collagen
Lgals1	Galectin-1	β-galactoside sugar binding

Dcn	Decorin	Small leucine-rich proteoglycan
Gpc4	Glypican 4	Heparin sulfate proteoglycan
Lum	Lumican	Small leucine-rich proteoglycan
Eln	Elastin	Elastic fibers
Col1a2	Collagen, type I, alpha 2 chain	Fibrillary collagen
Col6a2	Collagen, type VI, alpha 2 chain	Microfibrillar collagen
Col6a1	Collagen, type VI, alpha 1 chain	Microfibrillar collagen
Tgm3	Transglutaminase 3	Cross-links between glutamine and lysine
Postn	Periostin	Integrin ligand
Ogn	Osteoglycin / mimecan	Small leucine-rich proteoglycan
Prelp	Prolargin	Connects basement membranes to underlying connective tissue
Lama2	Laminin subunit alpha 2	Basement membrane glycoprotein
Lama5	Laminin subunit alpha 5	Basement membrane glycoprotein
Lamb2	Laminin subunit beta 2	Basement membrane glycoprotein
Muc2	Mucin-2	Mucus forming
Cspg4	Chondroitin sulfate proteoglycan 4	Chondroitin sulfate proteoglycan
Dpt	Dermatopontin	Small matricellular protein
Fbln5	Fibulin-5	Calcium-binding glycoprotein
Cluster 3	Name	Details
Fn1	Fibronectin	Fibrillar glycoprotein
Igfbp7	Insulin-like growth factor-binding protein 7	Controlling availability of IGFs
Tgfbi	Transforming growth factor, beta-induced	Binds collagen
Fbn1	Fibrillin-1	Microfibrillar glycoprotein
Lgals9	Galectin-9	β-galactoside sugar binding
Fbln1	Fibulin-1	Calcium-binding glycoprotein
Col4a2	Collagen, type IV, alpha 2 chain	Basement membrane
Pxdn	Peroxidasin	Extracellular heme-containing peroxidase
Plxnb2	Plexin-B2	Transmembrane receptor
Ecm1	Extracellular matrix protein 1	Soluble ECM protein
Itln1;Itln6	Intelectin-1	Immune lectin
Muc3	Mucin-3	Mucus forming
Hspg2	Perlecan	Basement membrane heparan sulfate proteoglycan
Dmbt1	Deleted in malignant brain tumors 1 protein	Glycoprotein
Vwf	Von Willebrand factor	Blood glycoprotein; hemostasis
Col18a1	Collagen, type XVIII, alpha 1 chain	Basement membrane; endostatin precursor
Anxa6	Annexin A6	Annexin
Itih1	Inter-alpha-trypsin inhibitor heavy chain H1	Serine-protease inhibitor
Ctsa	Cathepsin A	Serine protease
Agrn	Agrin	Proteoglycan
Adam10	A Disintegrin and metalloproteinase domain-containing protein 10	Sheds cell-surface molecules

Anxa3	Annexin A3	Annexin
Nid2	Nidogen-2	Basal lamina glycoprotein
Col4a1	Collagen, type IV, alpha 1 chain	Basement membrane
Anxa2	Annexin A2	Annexin
Nid1	Nidogen-1	Basal lamina glycoprotein
Ctsb	Cathepsin B	Serine protease
Lgals3	Galectin-3	β-galactoside sugar binding
Ctsd	Cathepsin D	Serine protease
Serpinh1	Serine protease inhibitor H1	Serine-protease inhibitor
Muc13	Mucin-13	Mucus forming
Tgm2	Transglutaminase 2	Cross-links between glutamine and lysine; autoantigen in celiac disease
Bgn	Biglycan	Small leucine-rich proteoglycan
Fbln2	Fibulin-2	Calcium-binding glycoprotein
Anxa5	Annexin A5	Annexin
Mfap2	Microfibrillar-associated protein 2	Antigen of elastin-associated microfibrils
St14	Suppressor of tumorigenicity 14	Serine protease
Anxa11	Annexin A11	Annexin
Anxa4	Annexin A4	Annexin
Loxl1	Lysyl oxidase homolog 1	Cross-linking Collagen and elastin
Lama4	Laminin subunit alpha 4	Basement membrane glycoprotein
Anxa7	Annexin A7	Annexin
P4ha1	Prolyl 4-hydroxylase subunit alpha-1	Collagen synthesis
Gpc1	Glypican 1	Heparin sulfate proteoglycan
Lman1	Lectin mannose-binding 1	Membrane protein; coagulation
Prg2	Proteoglycan 2	Eosinophil granule core component
Tnc	Tenascin C	Tenascin glycoprotein member; anti-adhesive
Efemp1	EGF-containing fibulin-like extracellular matrix protein 1	Calcium-binding glycoprotein
Vwa1	von Willebrand factor A domain containing 1	Promotes matrix assembly
Tinagl1	Tubulointerstitial nephritis antigen-like	Cysteine protease
Emilin1	Elastin microfibril interfacer 1	Glycoprotein; associated with elastin and microfibrils
Vwa5a	von Willebrand factor A domain containing 5A	Tumor suppressor
Serpinb1a	Serine protease inhibitor B1A	Serine-protease inhibitor
Plod3	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3	Lysyl hydroxylase

Table S4.

List of proteins in the three different clusters from **Fig. 5D** according to their abundance pattern across samples.

Gene	Direction	sequence
Mouse β-Actin	Forward	5- GGAGGGGGTTGAGGTGTT -3
Mouse β-Actin	Reverse	5- TGTGCACTTTTATTGGTCTCAAG -3
Mouse Fbn1	Forward	5- CGAGCCGCTAGTCCAGACA -3
Mouse Fbn1	Reverse	5- CTGTCCTGATGCAGAGAGGTC -3
Mouse Col18a1	Forward	5- CGAGAGGCACGGGCAATG -3
Mouse Col18a1	Reverse	5- GTCGTGCCGTGGAATAGGAG -3
Mouse Mmp2	Forward	5-CAGGGAATGAGTACTGGGTCTATT-3
Mouse Mmp2	Reverse	5-ACTCCAGTTAAAGGCAGCATCTAC-3
Mouse Mmp7	Forward	5-CTGCCACTGTCCCAGGAAG-3
Mouse Mmp7	Reverse	5-GGGAGAGTTTTCCAGTCATGG-3
Mouse Mmp9	Forward	5-GCAGAGGCATACTTGTACCG-3
Mouse Mmp9	Reverse	5-TGATGTTATGATGGTCCCACTTG-3
Mouse Mmp14	Forward	5-AGGCTGATTTGGCAACCATGA-3
Mouse Mmp14	Reverse	5-CCCACCTTAGGGGTGTAATTCTG-3
Mouse Timp1	Forward	5-CGAGACCACCTTATACCAGCG-3
Mouse Timp1	Reverse	5-ATGACTGGGGTGTAGGCGTA-3
Human HPRT	Forward	5- GCTATAAATTCTTTGCTGACCTGCTG -3
Human HPRT	Reverse	5- AATTACTTTTATGTCCCCTGTTGACTGG -3
Human UBC	Forward	5- GATTTGGGTCGCGGTTCTT -3
Human UBC	Reverse	5- TGCCTTGACATTCTCGATGGT -3
Human FBN1	Forward	5- GCGGAAATCAGTGTATTGTCCC -3
Human FBN1	Reverse	5- CAGTGTTGTATGGATCTGGAGC -3
Human COL18A1	Forward	5- GTTCCAGAGAATGCCGCTTG -3
Human COL18A1	Reverse	5- CCCATCTGAGTCATCGCCTT-3

Table S5.

Primer sequences used for quantitative real-time PCR analysis.

1290

Metabolite
Acetate
Alanine
Aspartate
Betaine
Choline
Ethanol
Formate
Glutamate
Glycine
Histidine
Isoleucine
Leucine
Methionine
Phenylalanine
Proline
Propionate
Putrescine
Succinate
Threonine
Tryptophan
Tyrosine
Valine
4-hydroxyl-L-proline

Table S6.

List of metabolites released from proteolytically cleaved ECMs.

Sample type	Group	Type of IBD	Sex	Age
Biopsy	Healthy	---	M	29
Biopsy	Healthy	---	M	74
Biopsy	Healthy	---	F	82
Biopsy	Healthy	---	F	55
Biopsy	Healthy	---	F	60
Biopsy	Healthy	---	M	45
Biopsy	Healthy	---	F	54
Biopsy	Healthy	---	F	67
Biopsy	Healthy	---	M	45
Biopsy	Healthy	---	M	50
Biopsy	Healthy	---	M	69
Surgical resection	IBD	UC	F	16
Surgical resection	IBD	Undetermined	M	43
Surgical resection	IBD	UC	F	36
Surgical resection	IBD	Crohn	M	48
Surgical resection	IBD	Crohn	M	38
Surgical resection	IBD	Undetermined	F	39
Biopsy	IBD	UC	F	26
Biopsy	IBD	UC	F	26
Biopsy	IBD	UC	M	53
Biopsy	IBD	UC	M	25

1295 **Table S7.**

List of human samples in each group of mRNA analysis displayed in **Fig. 7C**.