bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Title: Unraveling the state-specific nature of the native extracellular matrix via multidimensional characterization of its material properties

Authors: Elee Shimshoni[1], Ran Afik[1], Inna Solomonov[1], Idan Adir[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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Optimal exploitation of biological matrices for tissue engineering is hindered by incomplete understanding of their dynamic nature and the multidimensional complexity of their biomaterial properties. The unmet need in the field warrants investigation of the natural extracellular matrix as a complete and defined biomaterial. Here, we utilized two murine models to characterize native extracellular matrix alterations during inflammatory disease development. Biophysical and biochemical data integration allowed us to map discrete states unique in morphology, stiffness, and protein and metabolite compositions, which correlated with elevated matrix-degrading enzyme activity. Through the lens of this analysis we uncovered a silent pre-symptomatic tissue state, exhibited as a multidimensional biomaterial footprint, and matrix-derived clinically relevant biomarkers. Our holistic analysis showcases the native extracellular matrix as a dynamic and ultra-sensitive sensor of discrete tissue states. Successful matrix implementation in material research and therapeutics must be fine-tuned in a state-specific manner to ensure their desired bioactivity.

Keywords

Extracellular matrix, biomaterial, microscopy, proteomics, biomechanics

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Main

The extracellular matrix (ECM) is a natural biomaterial that 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], which is an intricate mesh of collagens, proteoglycans and glycoproteins and glucosaminoglycans (GAGs)[3]. For this reason, there is a growing interest in using ECM[4,5] or ECM-like materials as scaffolds for 3D cell cultures in regenerative medicine, cancer therapeutics and *in vitro* disease models[6–10]. To this end, hydrogels[11,12], artificial collagen gels[13], as well as matrix produced *in vitro* by fibroblasts[14] or sarcoma cells (Matrigel™)[15,16], have been extensively used and studied. Though each of these materials mimic certain properties of the native ECM – e.g., hydrogels for mechanical properties[17–19] and reconstituted cell-secreted ECM for specific components, they do not recapitulate the full complexity of the native ECM, which is mechanically[18], structurally[20] and compositionally[21] tissue-specific.

Previous studies have recognized extensive ECM remodeling events comparing healthy states to full-blown disease states, especially in cancer[22–25] and fibrosis[26–30], but limited information is available regarding inflammatory diseases. Nevertheless, changes in the expression or activity of specific ECM components or remodeling enzymes, has been demonstrated in inflammatory processes occurring in different tissues and under different conditions[31–37], suggesting that the real-life ECM is disease-state specific in inflammatory diseases as well. The main methods applied to gather information on the altered biomaterial properties of the ECM focus either on proteomic composition[21,38,39] or biomechanics[22,40,41].

These studies provide limited information about the ECM as a defined biomaterial as they often underscore only individual components or properties within the complex matrix. Accordingly, we have undertaken an integrative upgrade to the analysis ECM biomaterial

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

properties, which we propose is warranted. We chose to envision the ECM as a whole and analyze its biomaterial dynamics during inflammatory bowel diseases (IBD), characterized by intestinal inflammation, as a test case and probe for our approach.

IBD 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[31,42–54]. In light of this evidence, we have developed an ECM-centric approach, integrating a number of dimensions, for analysis of disease development and progression in intestinal inflammation. This highly sensitive paradigm for interrogation of ECM dynamics encompasses three dimension-metrics: 1) morphology 2) composition and 3) stiffness. These measurements have not been previously applied for the study of inflammatory diseases, and furthermore, have not been integrated into one comprehensive picture. This enabled us to meticulously define a clinically silent pathological tissue state occurring before the onset of inflammatory symptoms[55], which showcases a "pre-symptomatic ECM". This state is demonstrated both in mice at a pre-clinical time-point of an acute colitis model and in healthy IL-10 mice, which are prone to develop spontaneous chronic colitis, and is significantly different from that of healthy wild type (WT) mice. Importantly, as a consequence of analysis of this state using three dimension-metrics, we conclude that none of these approaches on its own would be sufficient to conclusively pinpoint and characterize this disease-prone state. As many types of inflammation in different tissues undergo a similar course of development, we anticipate that this type of characterization could be applied to uncover other ECM-based silent pathological states. In addition, as part of the ECM compositional analysis in the murine models, we pin pointed potential biomarkers. Remarkably, we were also able to identify them in tissues of IBD patients, thus demonstrating the translational potential of state-specific ECM analysis.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

This showcases the importance of the integrative native ECM-centric approach underscoring the ECM as an ultra-sensitive sensor of tissue states. Successful implementation of matrices in research and therapeutics must be fine-tuned in a state-specific manner to ensure their desired bioactivity.

Results

Characterization of murine colitis models for mapping natural matrix biomaterial properties and their dynamics

To probe the material properties of native ECM during disease-associated transitions we adopted established IBD murine models. Noteworthy, the study of IBD is almost exclusively immunological and cell-centric, with only limited study of the ECM and lack of information regarding its material properties, since it is relegated to background or contextual consideration. We have challenged this bystander role for the ECM[31].

To begin, we use a common murine model for IBD – the DSS-induced colitis model[56], which, with the appropriate calibration, results in a well-defined time-frame of acute colitis development in wild-type (WT) mice (**Fig. 1A**, see Methods for details). Disease progression was monitored by body weight measurement, endoscopic evaluation and histological analysis of colonic tissue, in accordance with commonly used clinically diagnostic methods[57] at different time points – day 0 (healthy WT), day 4 and day 10 (**Fig. 1** and **Fig. S1,** see Methods section for details on endoscopic scoring). Colon inflammation under these conditions peaks on day 10, as detected by endoscopy and histology (**Fig. S1A** and **Fig. 1B-D**), and by the immense increase in the amount of immune cells detected in the tissue via mass cytometry (**Fig. 1E-F**). 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. S1B**), appear healthy by

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

endoscopic evaluation, and do not display inflammation that is consistently detectable via histopathological analysis (**Fig. 1B** and **1D**). Therefore, we regard day 4 as the "pre-symptomatic" state. Despite showing no signs of inflammation by standard clinical tools[57], mice on day 4 display a statistically significant elevation (0.38% from of live cells compared to 0.021% in healthy WT) in the percentage of neutrophils in the colon compared to healthy WT, with no significant changes in other cell populations or in the total amount of immune cells (**Fig. 1E-F**). This finding demonstrates the earliest step of inflammation, albeit the relative number of infiltrating neutrophils is still very small.

Since our goal was to use a range of ECM states along the health-disease axis in intestinal inflammation, we also examined C57BL/6 IL-10$^{-/-}$ mice ("IL-10") that spontaneously develop chronic colon inflammation[58]. 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[59,60] (**Fig. 1A**, see Methods for details). IL-10 mice develop IBD-like symptoms, which can be detected and scored using colonoscopy similarly to the well-established DSS-induced model (**Fig. S1A**). Histopathology in this model shows prominent immune cell infiltration into the mucosa and has a different scoring system than DSS (see 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. 1D**). 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. 1C**). For comparing acute (day 10 of DSS) to chronic inflammation, we selected a stable chronic inflammation time point for further investigation (day

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

28), when mice appear to be at, or surpass, their initial weight before treatment (**Fig. S1B**). 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 were defined in two animal models: one for transient acute colitis (DSS) and the other for persistent chronic colitis (IL-10).

ECM morphological changes precede inflammatory symptoms

Having established pre-disease and disease state conditions, our first dimension of analysis was characterization of ECM 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 was achieved using scanning electron microscopy (SEM) on decellularized colon tissue from the same mice. Despite the downfalls of sample preparation, i.e., loss of certain components due to decellularization, EM provides us with high resolution images of the spatial organization of the ECM. SHG microscopy revealed that in the colon of the healthy WT mouse, the collagen (**Fig. 2A**) uniformly circumscribes the colonic crypts with little to no signal of fibrillar collagen between crypt borders (borders are pointed out by arrows in **Fig. S2A**). The tissue architecture in this state is well-organized, with a homogeneous crypt diameter (mean=53.7μm, average STD=10.8μm) and crypt wall thickness (mean=12.8μm, average STD=6.1μm) (**Fig. 2C**). High resolution SEM reveals that healthy crypts are uniformly covered by a dense mesh layer, which is most likely the basement membrane, according to its morphology[61] (**Fig. 2B**). Both imaging modalities reveal that acute inflammation (day 10 of DSS) has a profound effect on ECM structure. This state presents with destruction of tissue architecture (**Fig. 2A**), altered and heterogeneous crypt (mean=78.7μm, average STD=19.1μm) and wall size (mean=38.1μm,

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

average STD=19.4μm) (**Fig. 2C**), and disrupted ECM morphology (**Fig. 2B** and **Fig, S2B**), as well as regions containing the high-intensity fibrillar structures highlighted in **Fig. S2A**. These findings indicate that acute inflammation is characterized by regional ECM degradation and build up. Similarly, SHG imaging of chronic inflammation in IL-10 mice reveals deterioration of collagen structure supporting the crypt wall, with crypt boundaries defined by a lower intensity signal than its corresponding healthy state (**Fig. 2A**), and the presence of disoriented fibrillar structure between crypts (**Fig. S2A**). Crypt wall thickness, as well as their diameters, increases and collagen ultrastructure is heterogeneous compared to the healthy IL-10 mice, from a diameter of 60μm to 85.3μm and from a wall thickness of 20.1μm to 37.8μm (**Fig. 2C** and **Fig. S2C-D**). Further inspection of the ECM structure using SEM reinforced this finding: naïve IL-10 crypt wall ECM appears fibrous, whereas in chronically ill IL-10 mice, the crypt wall ECM appears disintegrated and perforated, with fibrillary structures of different morphology (**Fig. 2B** and **Fig. S2B**). Remarkably, at the pre-symptomatic state of the DSS model (day 4), thickening or deterioration of some crypt walls occurs, with changes in crypt diameter, along with the appearance of fibrillar structures in the spaces between crypts (**Fig. 2A** and **2C** and pointed out by arrows in **Fig. S2A**), indicating remodeling events of the colonic ECM. 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 basement membrane layer coating the interstitial ECM (**Fig. S2B**).

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 at high-resolution 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 and disrupted basement membrane, as pointed

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

out by arrows in **Fig. 2B**, and shown in more detail in **Fig. S2B**. 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, of ~90% for day 4 and ~75% for healthy IL-10 (**Fig. 2D**) compared to healthy WT (~6%). Thus, IL-10 mice, which are prone to developing chronic inflammation, develop a "leaky" basement membrane as a quasi-stable pre-symptomatic state, compared to the dense mesh covering the WT murine colon (**Fig. 2B** and **Fig. S2B**).

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 basement membrane, is one of the features of what we term the "pre-symptomatic ECM".

Disease onset is associated with a gradual loss of matrix rigidity

Having defined the structural and morphological differences of the portentous matrix in pre-symptomatic states, we applied a second dimension of our analysis, which interrogates the changes in the mechanical properties of the matrix. To do this, we used mesoscale atomic-force microscopy (AFM) on decellularized ECM scaffolds derived from colons of mice from both experimental models, using pre-symptomatic (day 4 of DSS and healthy IL-10) and pathological (day 10 of DSS and chronically ill IL-10) states, in addition to healthy WT.

We used micron-sized probes to apply pressure to a 500mmX500mm ECM scaffold. This permitted acquisition of the scaffolds' mesoscopically averaged, statistically robust Young's Modulus (YM) values[62,63], which describe the collective contributions of nanoscale molecular

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

components organized in micrometer-sized (70μmX70μm) domains in the ECM (see Methods for more details). Distributions of the measured YM values of colonic ECM samples are broad (**Fig. S3**), 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. S3** and **Table S3**) represents the elastic properties of a region of the ECM with lateral dimensions (and thickness) of several microns, which we term "mechanical domains". The median values and average median values of the YM of these mechanical domains are reported in **Fig. 2E-F**.

In all cases, we observe that the ECMs of inflamed colons are generally softer than those of healthy colons, as indicated by the median YM (**Fig. 2E-F** and **Table S2**). Importantly, the structurally modified ECM in the pre-symptomatic state (day 4) of the DSS model has mechanical domains with an intermediate rigidity (YM=9kPa), with values between those of healthy (YM=10.5kPa) and acutely inflamed (YM=6kPa) states of DSS. The ECM of the chronically ill IL-10 mice (YM=3kPa) is much softer than that of the healthy IL-10 mice (YM=11kPa) (**Fig. 2F** and **Table S2**). Remarkably, chronically ill IL-10 mice possess a YM value that is approximately two-fold lower than that of the acutely inflamed colon (**Fig. 2E-F** and **Table S2**).

Unexpectedly, we find that the rigidity of the ECM is similar in healthy IL-10 and healthy WT mice, both possessing a stiff colonic matrix, with median YM values of about 11kPa. This similarity is in contrast to clear structural differences between the two mice. These 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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 damage, considered a hallmark of disease progression. We find that loss of ECM rigidity is even more pronounced in the chronic colitis model than in severe acute disease.

Pre-symptomatic ECM changes coincide with epithelial gelatinase activity and neutrophils bearing remodeling enzymes

The gelatinases MMP-2 and MMP-9 primarily degrade basement membranes[64,65]. Therefore, having observed that the most prominent structural ECM alteration occurring at pre-symptomatic states involves perforation of the basement membrane underlying the epithelial cells of the colonic crypts and covering collagen fibrils, we set to examine the levels of ECM remodeling enzymes and 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 predominantly upregulated both at day 4 of the DSS model and in healthy IL-10 mice compared to healthy WT and the expression of TIMP-1, which inhibits MMP-9 and -2 was unchanged (**Fig. 3A**). Since the final regulatory step of MMPs occurs outside of the cell, with their activation, we used the activity-based assay, *in situ* zymography[66] to find the localization of active gelatinases in the tissue. Our analysis reveals that elevated (~1.5-2 fold) gelatinolytic activity is detected in the crypts of the pre-symptomatic tissues, compared to healthy WT (**Fig. 3B-C**).

In order to investigate the cellular sources of ECM remodeling enzymes at the transient day 4 state we used mass cytometry analysis that allows us to sort cells according to a large number of markers. This yields a comprehensive picture of the relative frequencies of different cell populations that express proteins of interest. The results reveal that epithelial cells are

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

prominent contributors of MMP-2, but the frequency of MMP-2-positive epithelial cells does not significantly change over the course of the disease (**Fig. S4A** and **S4D**).

Expectedly[47,67], day 10 samples have a large number of cells that stain positive for MMP-9 (**Fig. S4A**), but specific immune subsets, e.g., neutrophils and monocytes bear significant amounts of additional enzymes – lysyl oxidase (LOX), MMP-7,-13 and-14 (**Fig. S4B-C**). Interestingly, the only immune subset to begin infiltration on day 4, albeit at small quantities (0.38% of all cells, **Fig. 1F**), neutrophils, carry a variety of remodeling enzymes, including LOX, MMP-7,-9 and -13 (**Fig. 3D**). This indicates their contribution to the formation of the pre-symptomatic ECM.

This finding shows that while MMP-9, especially, is an established marker of active IBD[47,67], it may also display elevated activity in pre-clinical states, along with MMP-2. Together with the structural analysis in the previous section, we show that elevated ECM enzymatic degradation take part in the structural alterations in the ECM material that we observe at pre-symptomatic states.

Matrisome protein compositional dynamics reflects the state of the colon

To enhance the definition of the two distinct silent pre-symptomatic states, the third dimension of our analysis was the protein composition of the ECM biomaterial – the *matrisome*[21]. We used mass-spectrometry (MS)-based analysis of the tissue proteome, followed by statistical analysis and computational modeling. Eight to ten colon samples were analyzed from each of the five states: healthy WT, the two pre-symptomatic states (i.e., day 4 DSS, healthy IL-10), and two types of 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,

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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* [21].

We identified the ECM proteins that are significantly changed in abundance among states (see Methods for details). A scheme summarizing our results is shown in **Fig. 4A**, and in more detail in **Table S4**. 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**). This corresponds with some of the aforementioned structural changes – the newly apparent fibrillar structures appearing at day 4 may be correlated with the newly synthesized collagen type I, which can be cross-linked by transglutaminase, and the loss of mucin compromises the barrier function of the tissue, which is also aided by basement membrane perforation at these pre-symptomatic states. 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, which includes the reduction of several structural components (e.g., collagens, Laminin and Fibulin-5) (**Fig. 4C**), corresponding to the structural damage observed in ill states (**Fig. 2**). The two ill states, though clinically and functionally distinct, are also quite similar in terms of differentially abundant proteins, 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,

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

healthy IL-10 (**Table S4**), which is consistent with our finding that its rigidity is also the lowest of all states.

We also sought to test whether ECM remodeling leads to release of metabolites, as has been reported for signaling molecules[68–72], and how the repertoire of metabolites changes according to the unique ECM substrate of each specific tissue state. Subjecting the ECM of each of the states to degradation by the collagenase MMP-1, revealed that the pre-symptomatic states are similar to one another in their ECM-derived metabolite repertoires, and most importantly-distinct from the ill states and the healthy WT (**Fig. S5A** and **S5C**). Specifically, it appears that hydroxy-L-proline, which is a collagen-unique amino acid, is especially enriched among released metabolites from day 4 ECM (**Fig. S5B**). The distinct ECM-derived metabolite profiles released by the pre-symptomatic states add another parameter to their unique compositional ECM signature.

In summary, the compositional analysis of all experimental samples provides a molecular signature for the "pre-symptomatic ECM" in both acute and chronically ill models, which is compositionally distinct 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 the proteomic analysis, we have utilized machine-learning algorithms. Initially, we subjected the data to a supervised learning algorithm, a J48 pruned decision tree classifier[73], 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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. The resulting model is illustrated in **Fig. 4D**. 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[74,75]. The glycoprotein Fibrillin-1 (Fbn1), is part of the microfibril-forming Fibrillin family that is important for connective tissue integrity and function[76]; notably, mutations in the FBN1 gene lead to Marfan syndrome, which mostly manifests in skeletal, cardiovascular, and ocular abnormalities[77,78]. 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. 4D**). 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. 4E**). 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. 4F**), while Fbn1 expression 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 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. First, we used clustering to define groups of proteins that vary similarly across the experimental conditions. This allowed us to identify the dominant patterns in our high-dimensional dataset and avoid overfitting. We identified three clusters of proteins with similar abundance patterns (**Fig. 4G**). 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%). 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. Interestingly, the proteins in cluster 1, which are mostly protease 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 S5**). Both clusters reflect the tissue damage and ECM destruction characteristic of disease states. 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

between the ill and healthy states (**Fig. 4H**). 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 the ECM rigidity analysis, described above. PC1 values of chronic illness reflect a 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.

Human IBD recapitulates murine matrisomic dynamics and validate biomarkers

In order to assess to clinical relevance of our findings, 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[79]. 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. 5A**). 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 healthy donors, and vice versa (**Fig. 5B**).

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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

XVIII and Fibrillin-1 are significantly upregulated, in concordance with the decision tree applied to the proteomic data from our murine colitis samples (**Fig. 5C**). 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 control with little to no staining (**Fig. 5D-E**).

Taken together, analysis of published human proteomic datasets, qPCR and immunostaining of clinical biopsies, validate our findings. We confirmed the predictive value of biomarkers Collagen XVIII and Fibrillin-1, which we identified using a machine learning classification analysis. Proteomics with machine learning classification therefore represents a powerful application of one of the components of our analysis. This sort of investigation may have clinical diagnostic applications for inflammatory diseases.

Discussion

In this paper, we revealed the sensitivity of the ECM to the subtleties of different stages of the pathological process via an integrated multidimensional characterization of the native ECM's material properties, which has with both bioengineering and clinical relevance. Our findings underscore the heterogeneous and dynamic nature of the ECM that renders it not only tissue-specific, but also state-specific. As a test case, we used IBD, which has a number of immunologically well-studied mouse models, allowing us to characterize holistic and multidimensional ECM dynamics throughout inflammatory disease development and progression. Our results provide a high-definition representation of state-specific ECM signatures, both in healthy and in acute and chronic inflammation, but most strikingly, we uncovered a silent pathology detectable even before the onset of clinical symptoms (**Scheme 1**).

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A number of interesting observations stem from unraveling the state-specificity of the ECM, expressed by the changes in each biomaterial property on its own, and most importantly – in the comprehensive picture formed by their integration.

First, our analysis uncovered what we term a "pre-symptomatic ECM", a reproducible signature that is virtually identical in structure, proteomic and ECM-derived metabolite compositions, in acute and chronic models, and crucially, is distinct from that of the healthy WT mouse. Our analysis exposed several aspects and processes of remodeling that most likely contribute to the formation of this signature. This includes basement membrane degradation (**Fig. 2**), fibrillary collagen deposition (**Fig. 2** and **Fig. 4A**), differential proteolytic activity of gelatinases in the epithelium, as well as a small number of infiltrating neutrophils bearing ECM remodelers (**Fig. 3**). We propose that the disruption of the basement membrane underlying the epithelium (**Fig. 2B**), via proteolytic degradation (**Fig. 3**) by the epithelium and a small number of newly arriving neutrophils, allows access to immune cell infiltration. Though the results support MMP-9 as a therapeutic target in IBD[48],our data also indicates that the elevated gelatinolytic activity apparent in the epithelium (**Fig. 3B**), can be attributed to the activation and expression levels of MMP-2 per cell (**Fig. S4A** and **S4D**), and not to changes in the number of MMP-2-positive cells. Importantly, both the epithelium and the basement membrane are key players in the barrier function of all mucosal tissues[37], providing separation between the external microbial environment and the immune system. In addition, neutrophils bearing MMP-9 are often the first immune cell type to initiate the inflammatory process[80]. Hence, our data is in line with the knowledge of mucosal tissue function, strongly suggesting that a similar pre-symptomatic state would exist in other diseases.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Elevated gelatinolytic activity at pre-symptomatic states (**Fig. 3**) is correlated with elevated ECM degradation, which is supported by the prominent accumulation of the collagen-unique amino acid, hydroxyproline[81] that is proteolytically released from day 4 ECM (**Fig. S5B**). These findings significantly expand the understanding of how ECM dynamics embody and participate in the early stages of the inflammatory process. Moreover, they expand the hypothesis raised by a previous finding that defined only a single early ECM remodeling feature at early stages of the DSS model – changes in the deposition of the abundant ECM glycosaminoglycan component, hyaluronic acid[82].

The second discrete feature highlighted by our characterization 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 – indicating either ill or pre-symptomatic states, respectively. 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[83]. 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 ECM feature may have similar prognostic impact. Thus, machine learning as part of our compositional analysis has already demonstrated its potential for identifying relevant clinical markers and disease prognostic indicators.

A third advance of our analysis 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

analysis (**Fig. 4H**) differentiates between the five states, and moreover, orders the five states in complete correspondence to other dimensions of analysis. This is especially 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 (**Fig. 2**). For example, the IL-10 mouse exhibits less structural damage than the WT mouse state in the same group in terms of the preserved, albeit altered, crypt macrostructure, and correspondingly the PC1 score is higher. Hence, our comprehensive analysis addresses the fundamental need for a means for mapping of higher order structural organization to protein amounts, measured by quantitative methods.

Finally, synthesizing the data generated by the analysis allows us to differentiate acute and chronic disease, and not only to grossly divide the samples into three groups that we see at first glance. In our model, the duration and difference in tissue resilience of each disease type (i.e., acute vs. chronic) explains the position of the 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 reflects the process of colitis management by the tissue to maximize its functionality in light of ongoing challenge. This is suggested by the regular, though altered, crypt pattern. Finally, to support tissue resilience and perhaps promote maintenance of inflammation, long-term chronic inflammation allows more time for ECM reorganization, such as the observed changes in differentially abundant proteins (**Fig. 4A**) and

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

mechanical changes (**Fig. 2F**). 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 interpretation of our findings, demonstrating the response of gut tissue to acute and chronic disease, both in ill and pre-symptomatic states, highlights the ability of the ECM, as a tissue-spanning biomaterial, to sense subtle variations in tissue state.

In summary, our integrative approach permits a well-rounded picture, encompassing multiple levels of detail, of the ECM – a complex and substantial biomaterial that comprises tissues. We conclude that by manifesting the pathophysiological state of a tissue, the ECM reveals silent disease-prone states that occur before clinical symptoms. The lessons learned from the observations we made in intestinal inflammation can be applied to other scenarios of inflammation due to the similarities of its onset in other mucosal tissues. The growing interest in mimicking ECM scaffolds for basic and translational research requires a more detailed, yet holistic, understanding of tissue and state-specific ECM signatures. We anticipate that this work will inspire the design of novel physiological scaffolds, as well as fuel mechanistic studies on the means by which early morphological and biomechanical events leave the tissue vulnerable or prone to other ECM/tissue pathologies such as cancer, fibrosis and inflammation in other organs.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Acknowledgments

We thank Dr. Ori Brenner for assistance in histopathological scoring of intestinal inflammation. Also, 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, IA, AS, MA, LP and VG; Formal Analysis, ES, IA, 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, ISo, IS; Visualization, ES, MA, LP, VG and AP; Supervision, ISo, CV, DSS, AP, PT, TG, PM, CL, UA and IS; Funding Acquisition, IS, AP, TG, PM.

Declaration of Interests

The authors declare no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Frantz, C., Stewart, K. M. & Weaver, V. M. The extracellular matrix at a glance. *J. Cell Sci.* **123**, 4195–200 (2010).
2. Bosman, F. T. & Stamenkovic, I. Functional structure and composition of the extracellular matrix. *J. Pathol.* **200**, 423–8 (2003).
3. Hynes, R. O. The extracellular matrix: not just pretty fibrils. *Science* **326**, 1216–1219 (2009).
4. Badylak, S. F. The extracellular matrix as a biologic scaffold material. *Biomaterials* **28**, 3587–3593 (2007).
5. Badylak, S. F., Freytes, D. O. & Gilbert, T. W. Extracellular matrix as a biological scaffold material: Structure and function. *Acta Biomater.* **23**, S17–S26 (2015).
6. Gu, L. & Mooney, D. J. Biomaterials and emerging anticancer therapeutics: engineering the microenvironment. *Nat. Rev. Cancer* **16**, 56–66 (2016).
7. Hinderer, S., Layland, S. L. & Schenke-Layland, K. ECM and ECM-like materials — Biomaterials for applications in regenerative medicine and cancer therapy. *Adv. Drug Deliv. Rev.* **97**, 260–269 (2016).
8. Almouemen, N., Kelly, H. M. & O'Leary, C. Tissue Engineering: Understanding the Role of Biomaterials and Biophysical Forces on Cell Functionality Through Computational and Structural Biotechnology Analytical Methods. *Comput. Struct. Biotechnol. J.* **17**, 591–598 (2019).
9. Wade, R. J. & Burdick, J. A. Engineering ECM signals into biomaterials. *Mater. Today* **15**, 454–459 (2012).
10. Chew, S. A. & Danti, S. Biomaterial-Based Implantable Devices for Cancer Therapy. *Adv. Healthc. Mater.* **6**, 1600766 (2017).
11. Lee, K. Y. & Mooney, D. J. Hydrogels for tissue engineering. *Chem. Rev.* **101**, 1869–79 (2001).
12. Khan, F. & Tanaka, M. Designing Smart Biomaterials for Tissue Engineering. *Int. J. Mol. Sci.* **19**, (2018).
13. Brodkin, K. R., García, A. J. & Levenston, M. E. Chondrocyte phenotypes on different extracellular matrix monolayers. *Biomaterials* **25**, 5929–5938 (2004).
14. Magnan, L. *et al.* Characterization of a Cell-Assembled extracellular Matrix and the effect of the devitalization process. *Acta Biomater.* **82**, 56–67 (2018).
15. Hughes, C. S., Postovit, L. M. & Lajoie, G. A. Matrigel: A complex protein mixture required for optimal growth of cell culture. *Proteomics* **10**, 1886–1890 (2010).
16. Benton, G., George, J., Kleinman, H. K. & Arnaoutova, I. P. Advancing science and technology via 3D culture on basement membrane matrix. *J. Cell. Physiol.* **221**, 18–25 (2009).
17. Chaudhuri, O. *et al.* Hydrogels with tunable stress relaxation regulate stem cell fate and activity. *Nat. Mater.* **15**, 326–34 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

18. Engler, A. J., Sen, S., Sweeney, H. L. & Discher, D. E. Matrix Elasticity Directs Stem Cell Lineage Specification. *Cell* **126**, 677–689 (2006).

19. Cameron, A. R., Frith, J. E. & Cooper-White, J. J. The influence of substrate creep on mesenchymal stem cell behaviour and phenotype. *Biomaterials* **32**, 5979–5993 (2011).

20. Brown, B. N. *et al.* Surface characterization of extracellular matrix scaffolds. *Biomaterials* **31**, 428–37 (2010).

21. Naba, A. *et al.* The Matrisome: In Silico Definition and In Vivo Characterization by Proteomics of Normal and Tumor Extracellular Matrices. *Molecular & Cellular Proteomics* **11**, M111.014647-M111.014647 (2012).

22. Brauchle, E. *et al.* Biomechanical and biomolecular characterization of extracellular matrix structures in human colon carcinomas. *Matrix Biol.* **68–69**, 180–193 (2018).

23. Grossman, M. *et al.* Tumor cell invasion can be blocked by modulators of collagen fibril alignment that control assembly of the extracellular matrix. *Cancer Res.* (2016).

24. Afik, R. *et al.* Tumor macrophages are pivotal constructors of tumor collagenous matrix. *J. Exp. Med.* jem.20151193 (2016). doi:10.1084/jem.20151193

25. Lu, P., Weaver, V. M. & Werb, Z. The extracellular matrix: A dynamic niche in cancer progression. *J. Cell Biol.* **196**, 395–406 (2012).

26. Georges, P. C. *et al.* Increased stiffness of the rat liver precedes matrix deposition: implications for fibrosis. *Am. J. Physiol. Gastrointest. Liver Physiol.* **293**, G1147-54 (2007).

27. Bonnans, C., Chou, J. & Werb, Z. Remodelling the extracellular matrix in development and disease. *Nat. Rev. Mol. Cell Biol.* **15**, 786–801 (2014).

28. Cox, T. R. & Erler, J. T. Remodeling and homeostasis of the extracellular matrix: implications for fibrotic diseases and cancer. *Dis. Model. Mech.* **4**, 165–78 (2011).

29. Herrera, J., Henke, C. A. & Bitterman, P. B. Extracellular matrix as a driver of progressive fibrosis. *J. Clin. Invest.* **128**, 45–53 (2018).

30. Upagupta, C., Shimbori, C., Alsilmi, R. & Kolb, M. Matrix abnormalities in pulmonary fibrosis. *Eur. Respir. Rev.* **27**, 180033 (2018).

31. 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* **64**, 367–72 (2015).

32. 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* **280**, 2165–2179 (2013).

33. Wu, F. *et al.* A novel role of the lumican core protein in bacterial lipopolysaccharide-induced innate immune response. *J. Biol. Chem.* **282**, 26409–17 (2007).

34. Sorokin, L. The impact of the extracellular matrix on inflammation. *Nat. Rev. Immunol.* **10**, 712–723 (2010).

35. de la Motte, C. A. Hyaluronan in intestinal homeostasis and inflammation: implications

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

for fibrosis. *Am. J. Gastrointest. Liver Physiol.* **301**, 945–949 (2011).

36. Khokha, R., Murthy, A. & Weiss, A. Metalloproteinases and their natural inhibitors in inflammation and immunity. *Nat. Rev. Immunol.* **13**, 649–665 (2013).

37. Talmi-Frank, D. *et al.* Extracellular Matrix Proteolysis by MT1-MMP Contributes to Influenza-Related Tissue Damage and Mortality. *Cell Host Microbe* **20**, 458–470 (2016).

38. Naba, A. *et al.* Characterization of the Extracellular Matrix of Normal and Diseased Tissues Using Proteomics. *J. Proteome Res.* **16**, 3083–3091 (2017).

39. Massey, V. L. *et al.* The hepatic "matrisome" responds dynamically to injury: Characterization of transitional changes to the extracellular matrix in mice. *Hepatology* **65**, 969–982 (2017).

40. Mambetsariev, I. *et al.* Stiffness-Activated GEF-H1 Expression Exacerbates LPS-Induced Lung Inflammation. *PLoS One* **9**, e92670 (2014).

41. Schedin, P. & Keely, P. J. Mammary gland ECM remodeling, stiffness, and mechanosignaling in normal development and tumor progression. *Cold Spring Harb. Perspect. Biol.* **3**, a003228 (2011).

42. Bailey, C. J. *et al.* Distribution of the matrix metalloproteinases stromelysin, gelatinases A and B, and collagenase in Crohn's disease and normal intestine. *J. Clin. Pathol.* **47**, 113–116 (1994).

43. Pender, S. L. F. & MacDonald, T. T. Matrix metalloproteinases and the gut - New roles for old enzymes. *Current Opinion in Pharmacology* **4**, 546–550 (2004).

44. Meijer, M. J. W. *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.* **39**, 733–9 (2007).

45. Gao, Q. *et al.* Expression of matrix metalloproteinases-2 and -9 in intestinal tissue of patients with inflammatory bowel diseases. *Dig. Liver Dis.* **37**, 584–92 (2005).

46. Garg, P. *et al.* Selective ablation of matrix metalloproteinase-2 exacerbates experimental colitis: contrasting role of gelatinases in the pathogenesis of colitis. *J. Immunol.* **177**, 4103–12 (2006).

47. Baugh, M. D. *et al.* Matrix metalloproteinase levels are elevated in inflammatory bowel disease. *Gastroenterology* **117**, 814–822 (1999).

48. Sela-Passwell, N. *et al.* Antibodies targeting the catalytic zinc complex of activated matrix metalloproteinases show therapeutic potential. *Nat. Med.* **18**, 143–7 (2012).

49. Castaneda, F. E. *et al.* Targeted deletion of metalloproteinase 9 attenuates experimental colitis in mice: central role of epithelial-derived MMP. *Gastroenterology* **129**, 1991–2008 (2005).

50. Salmela, M. T. *et al.* 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.* **39**, 1095–104 (2004).

51. Hermano, E., Lerner, I. & Elkin, M. Heparanase enzyme in chronic inflammatory bowel disease and colon cancer. *Cell. Mol. Life Sci.* **69**, 2501–13 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

52. Waterman, M. *et al.* Heparanase upregulation by colonic epithelium in inflammatory bowel disease. *Mod. Pathol.* **20**, 8–14 (2007).

53. Lohr, K. *et al.* Extracellular matrix protein lumican regulates inflammation in a mouse model of colitis. *Inflamm. Bowel Dis.* **18**, 143–151 (2012).

54. 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.* **155**, 257–265 (2009).

55. Stange, E. F. *et al.* European evidence-based Consensus on the diagnosis and management of ulcerative colitis: Definitions and diagnosis. *J. Crohns. Colitis* **2**, 1–23 (2008).

56. Okayasu, I. *et al.* A novel method in the induction of reliable experimental acute and chronic ulcerative colitis in mice. *Gastroenterology* **98**, 694–702 (1990).

57. Stange, E. F. *et al.* European evidence based consensus on the diagnosis and management of Crohn's disease: definitions and diagnosis. *Gut* **55 Suppl 1**, i1-15 (2006).

58. Kühn, R., Löhler, J., Rennick, D., Rajewsky, K. & Müller, W. Interleukin-10-deficient mice develop chronic enterocolitis. *Cell* **75**, 263–274 (1993).

59. Glocker, E.-O. *et al.* Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N. Engl. J. Med.* **361**, 2033–45 (2009).

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

61. van der Rest, M. & Garrone, R. Collagen family of proteins. *FASEB J.* (1991).

62. Indrieri, M., Podestà, A., Bongiorno, G., Marchesi, D. & Milani, P. Adhesive-free colloidal probes for nanoscale force measurements: production and characterization. *Rev. Sci. Instrum.* **82**, 023708 (2011).

63. 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.* **86**, 033705 (2015).

64. Van den Steen, P. E. *et al.* Biochemistry and Molecular Biology of Gelatinase B or Matrix Metalloproteinase-9 (MMP-9). *Crit. Rev. Biochem. Mol. Biol.* **37**, 375–536 (2002).

65. Zeng, Z.-S., Cohen, A. M. & Guillem, J. G. 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* **20**, 749–755 (1999).

66. George, S. J. & Johnson, J. L. In Situ Zymography. in 271–277 (Humana Press, Totowa, NJ, 2010). doi:10.1007/978-1-60327-299-5_17

67. 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.* **57**, 706–712 (2012).

68. Hashimoto, G. *et al.* Matrix metalloproteinases cleave connective tissue growth factor and reactivate angiogenic activity of vascular endothelial growth factor 165. *J. Biol. Chem.* **277**, 36288–95 (2002).

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

69. Whitelock, J. M., Murdoch, A. D., Iozzo, R. V & Underwood, P. A. 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.* **271**, 10079–86 (1996).

70. Mañes, S. *et al.* The matrix metalloproteinase-9 regulates the insulin-like growth factor-triggered autocrine response in DU-145 carcinoma cells. *J. Biol. Chem.* **274**, 6935–45 (1999).

71. 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.* **314**, 2501–2514 (2008).

72. 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.* **41**, 233–264 (2004).

73. Salzberg, S. L. C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Mach. Learn.* **16**, 235–240 (1994).

74. Marneros, A. G. & Olsen, B. R. Physiological role of collagen XVIII and endostatin. *FASEB J.* **19**, 716–28 (2005).

75. Utriainen, A. *et al.* Structurally altered basement membranes and hydrocephalus in a type XVIII collagen deficient mouse line. *Hum. Mol. Genet.* **13**, 2089–2099 (2004).

76. Sakai, L. Y., Keene, D. R. & Engvall, E. Fibrillin, a new 350-kD glycoprotein, is a component of extracellular microfibrils. *J. Cell Biol.* **103**, 2499–509 (1986).

77. Dietz, H. C. *et al.* Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. *Nature* **352**, 337–339 (1991).

78. Hollister, D., Godfrey, M., Sakai, L. & Pyeritz, R. Immunohistologic Abnormalities of the Microfibrillar-Fiber System in the Marfan Syndrome. *N. Engl. J. Med.* **323**, 152–9 (1990).

79. Bennike, T. B. *et al.* Neutrophil Extracellular Traps in Ulcerative Colitis: A Proteome Analysis of Intestinal Biopsies. *Inflamm. Bowel Dis.* **21**, 2052–67 (2015).

80. Bardoel, B. W., Kenny, E. F., Sollberger, G. & Zychlinsky, A. The Balancing Act of Neutrophils. *Cell Host Microbe* **15**, 526–536 (2014).

81. Udenfriend, S. Formation of hydroxyproline in collagen. *Science* **152**, 1335–1340 (1966).

82. Kessler, S. *et al.* Hyaluronan (HA) deposition precedes and promotes leukocyte recruitment in intestinal inflammation. *Clin. Transl. Sci.* **1**, 57–61 (2008).

83. Kugathasan, S. *et al.* Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. *Lancet* **389**, 1710–1718 (2017).

84. Becker, C., Fantini, M. C. & Neurath, M. F. High resolution colonoscopy in live mice. *Nat. Protoc.* **1**, 2900–2904 (2006).

85. Zigmond, E. *et al.* Ly6C hi Monocytes in the Inflamed Colon Give Rise to Proinflammatory Effector Cells and Migratory Antigen-Presenting Cells. *Immunity* (2012). doi:10.1016/j.immuni.2012.08.026

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

86. Behbehani, G. K. *et al.* Transient partial permeabilization with saponin enables cellular barcoding prior to surface marker staining. *Cytometry. A* **85**, 1011 (2014).

87. Zunder, E. R. *et al.* Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. *Nat. Protoc.* **10**, 316–333 (2015).

88. 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* **100**, 2507–16 (2012).

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

90. Butt, H.-J., Cappella, B. & Kappl, M. Force measurements with the atomic force microscope: Technique, interpretation and applications. *Surf. Sci. Rep.* **59**, 1–152 (2005).

91. Schillers, H. *et al.* Standardized Nanomechanical Atomic Force Microscopy Procedure (SNAP) for Measuring Soft and Biological Samples. *Sci. Rep.* **7**, 5117 (2017).

92. Cramér, H. Mathematical Methods of Statistics. *Princeton University Press* (1999).

93. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

94. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry – based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).

95. Cox, J. *et al.* Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).

96. Cox, J. *et al.* Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).

97. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(- Delta Delta Ct) Method. *Methods* **25**, 402–408 (2001).

98. 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.* 4–6 (2013).

99. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).

100. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci.* **98**, 5116–5121 (2001).

101. Solomonov, I. *et al.* Distinct biological events generated by ECM proteolysis by two homologous collagenases. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 10884–9 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

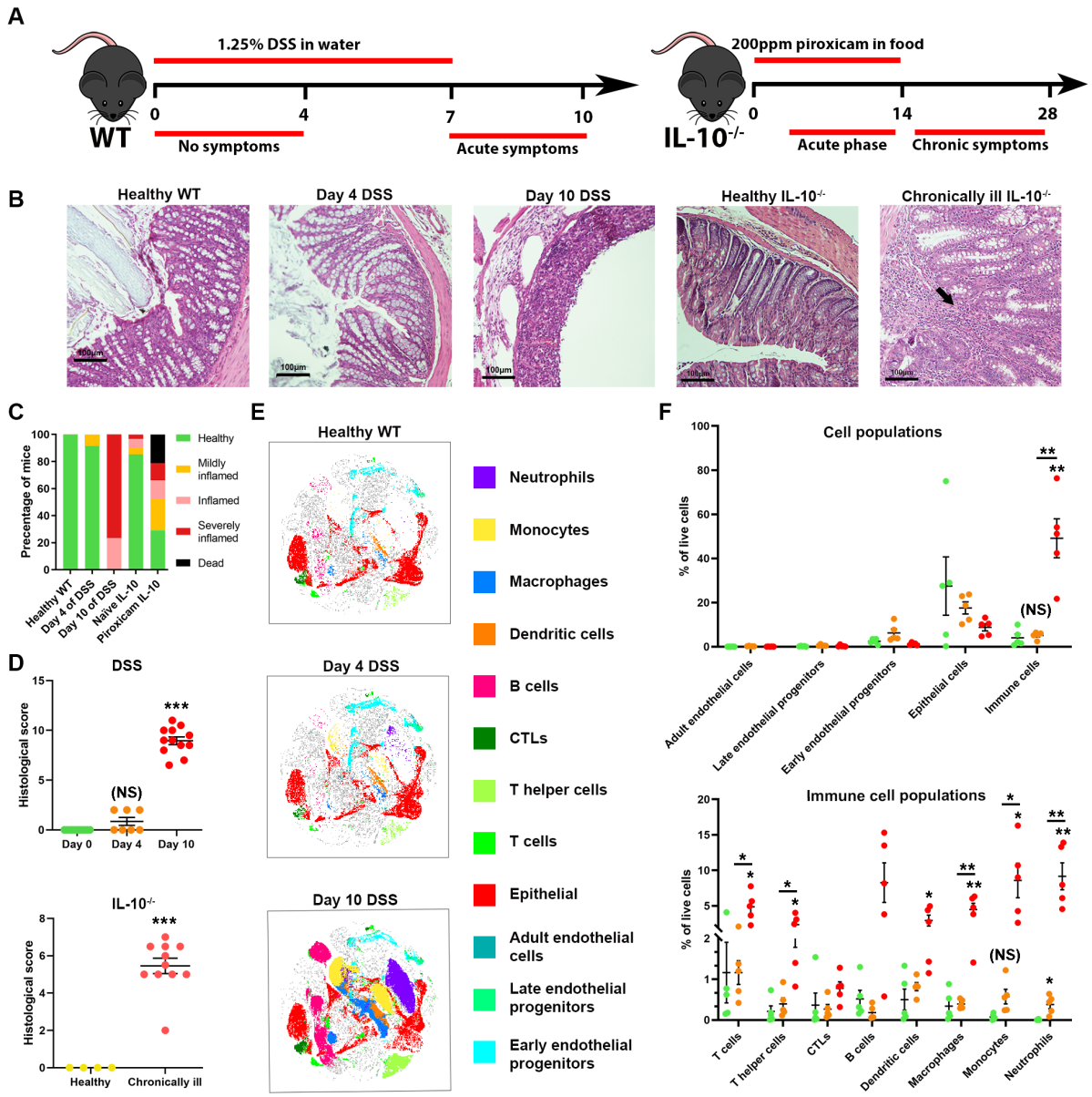


Fig.1 Definition of the disease course and tissue states in the two murine colitis models

(A) 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. (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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

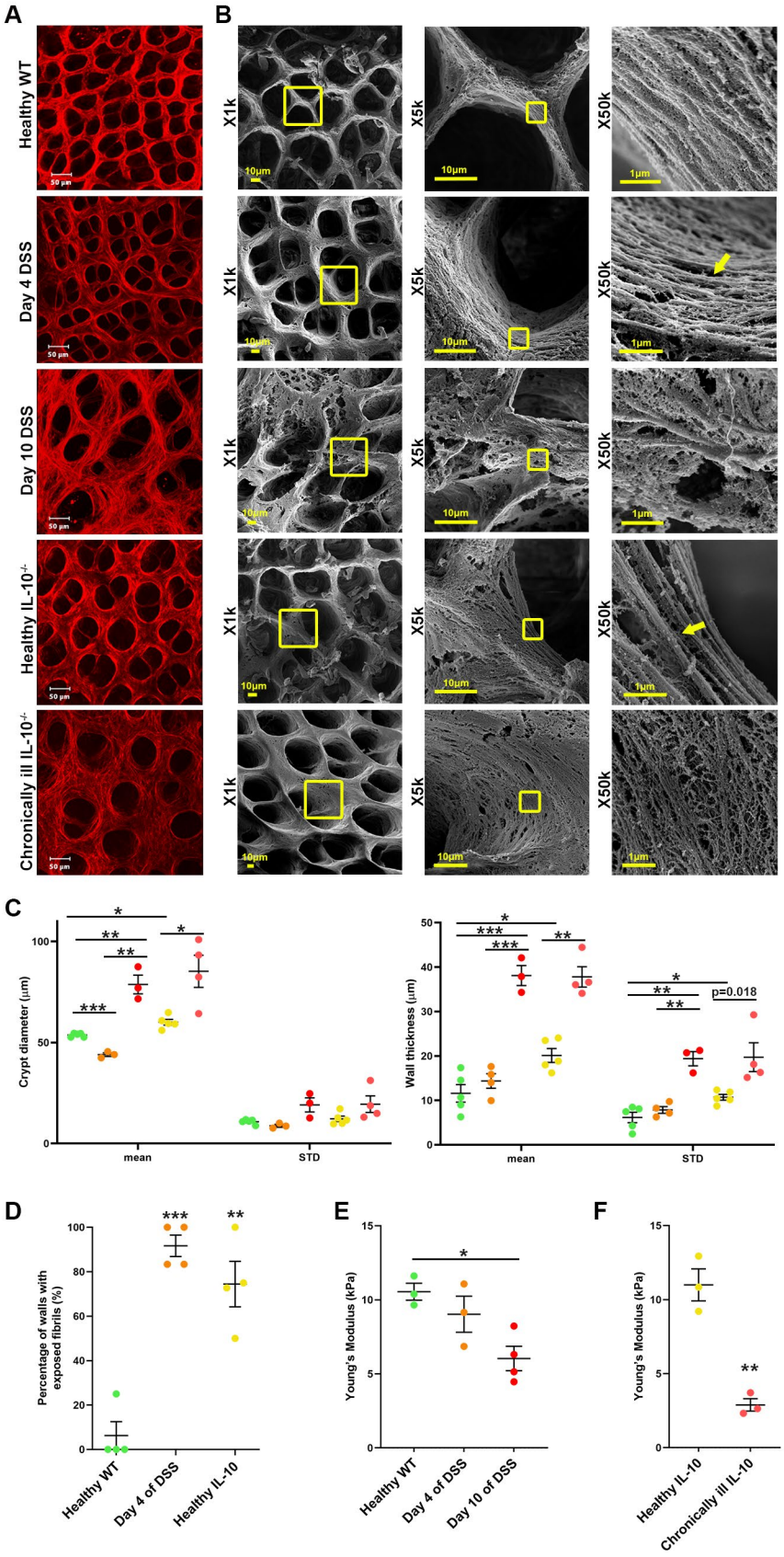
of the acute model, but is substantial on day 10. In the PAC IL-10$^{-/-}$ model large amounts of immune cell infiltrate is observed in the mucosa (indicated by arrow), but is absent in the healthy IL-10$^{-/-}$ mouse. (C) Quantification of the percentage of mice in each clinical category according to endoscopic evaluation (0-4: "healthy"; 5-7 "mildly inflamed"; 8-11 "inflamed"; and 12-15 "severely inflamed", see **Fig. S1A** and Methods for details) 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. Only mice evaluated as "healthy" were used for further analysis as the day 4 of DSS or healthy IL-10$^{-/-}$ states. (D) Histological scoring of H&E stained sections, such as the ones shown in B. Details on scoring are available in the Methods. 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$^{-/-}$ mice display a significant increase in histopathological score. The plot displays the scores of individual animals and the bars represent the state's mean±s.e.m. Statistical significance was determined using a student's t-test with a Bonferroni correction for two comparisons to healthy WT for the DSS model (NS: P> 0.025, *** P<0.0005) and with no correction, compared to healthy IL-10$^{-/-}$ in the PAC IL-10$^{-/-}$ model (*** P<0.001). (E) tSNE analysis of mass cytometry demonstrating the changes in cell composition at each time point of the DSS acute colitis model. Each plot represents a pool of all five animals analyzed per state and each point corresponds to one cell. (F) Scatter dot plots based on the mass cytometry results depicting the changes in cell population relative frequencies (out of all live cells) at each state: healthy WT (green), day 4 of DSS (orange), day 10 of DSS (red). Note that the only statistically significant change in cell composition on day 4 is a slight increase in the relative requency of

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

neutrophils. Number of animals: n=5 per state; Bonferroni correction for three comparisons:

780 *P<0.0167, **P<0.00333.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 2 The ECM suffers structural damage and softening 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 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 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 and crypt-wall thickness in second-harmonic images on all five states – healthy WT (green), DSS day 4 (orange), DSS day 10 (red), healthy IL-10$^{-/-}$ (yellow) and ill IL-10$^{-/-}$ (pink). In oval crypts, the largest diameter was always chosen, and the shortest distance was taken for wall thickness measurements. Note that each state has a different mean diameter and wall thickness. Changes in the standard deviations (STDs) of these measures per animal are also apparent, demonstrating their heterogeneity. Both of these measures indicate the architectural changes taking place in 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

represent the mean±s.e.m of the state, while each dot corresponds to the mean/standard deviation of one animal, which is based on the analysis of 2-8 440µmX440µm images capturing n>30 crypts per animal. One outlier was removed in the day 4 group. (D) The relative 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+F) 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. S3**. (E) 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. (F) The results show a significant reduction in ECM rigidity between the healthy and chronically inflamed states, as denoted by the asterisks (**P<0.01).

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

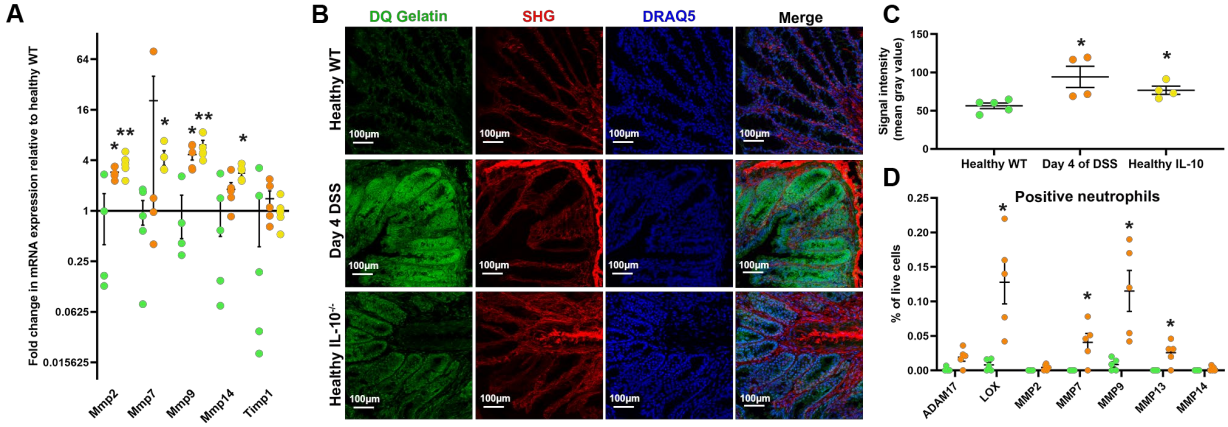
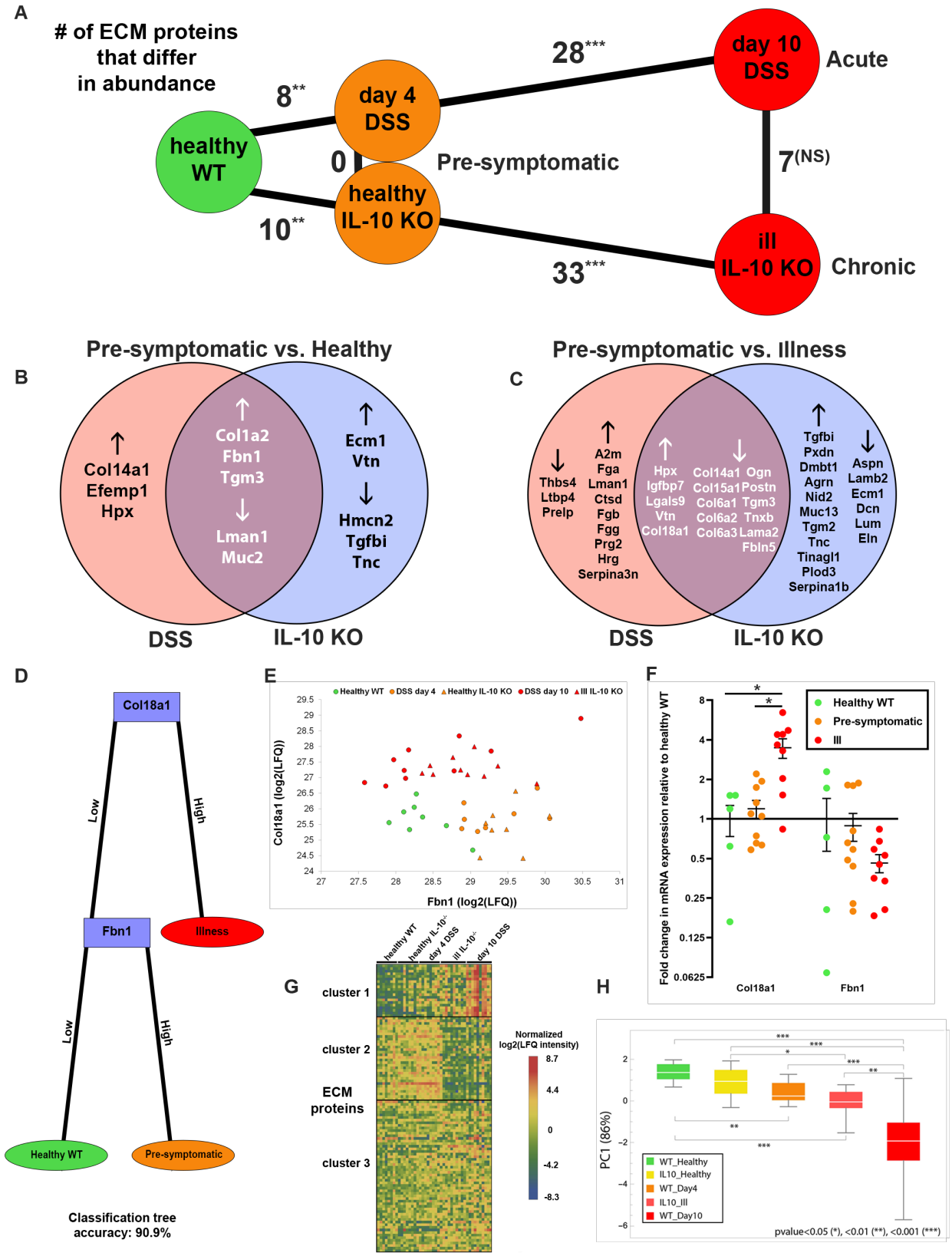


Fig. 3 ECM remodeling is enhanced in pre-symptomatic states by different cell types

(A) 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. (B) *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). (C) Quantification of signal intensity in the mucosa, revealing that the elevation in gelatinase activity is statistically significant. (D) Scatter dot plots based on the mass cytometry analysis depicting the changes in the relative frequencies (out of all live cells) of neutrophil populations positive for the indicated remodeling enzymes at each state: healthy WT (green), day 4 of DSS (orange). Additional data (including neutrophils on day 10) from this analysis appears in **Fig. S4**. Note that a significant amount of infiltrating neutrophils carry ECM remodelers. Number of animals: n=5 per state; Bonferroni correction for three comparisons: *P<0.0167, **P<0.00333.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



840

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 4 ECM proteomics and machine-learning-based analyses decode compositional differences and similarities during colitis and unravels potentially pivotal components

(A) Scheme showing the number of differentially abundant ECM proteins, between pairs in the five different states, 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+C) Venn diagrams 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) in B; and 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) in C. The direction of change in abundance in pre-symptomatic states is indicated by arrows. Note the differentially abundant proteins shared among the two comparisons (purple) in both diagrams. The intersection between the differentially abundant proteins in both comparisons is statistically significant ($p=1.0x10^{-4}$ in B and $p=1.2x10^{-5}$ in C). Statistical significance in this figure was calculated by the sum of hypergeometric probabilities for achieving each number or above it. (D) 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[73] 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

highest accuracy (90.9%) of all proteins. Col18a1 threshold: 26.67 log2(LFQ intensity). Fbn1 threshold: 28.76 log2(LFQ intensity). (E) 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. (F) 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 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. (G) 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 clinical state is indicated on the x axis. Protein abundances were normalized to the mean of each protein. (H) 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. *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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

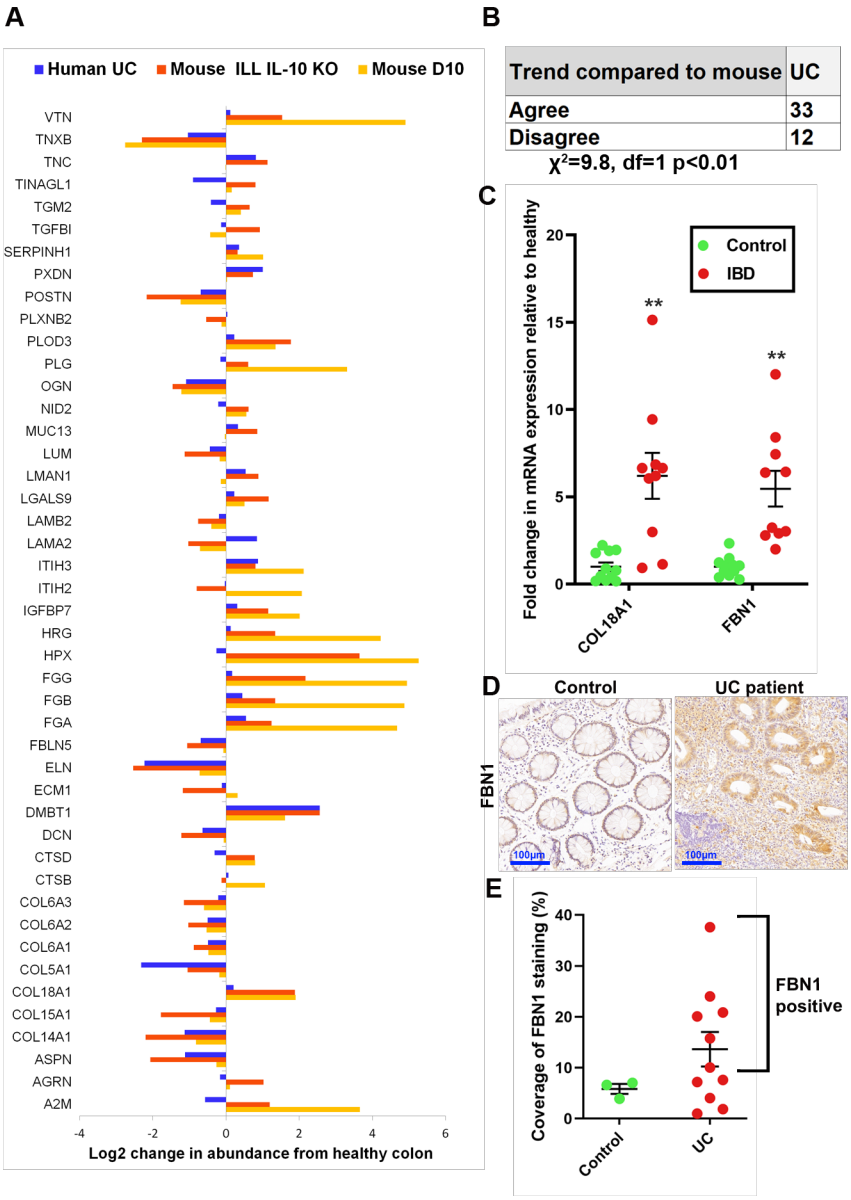


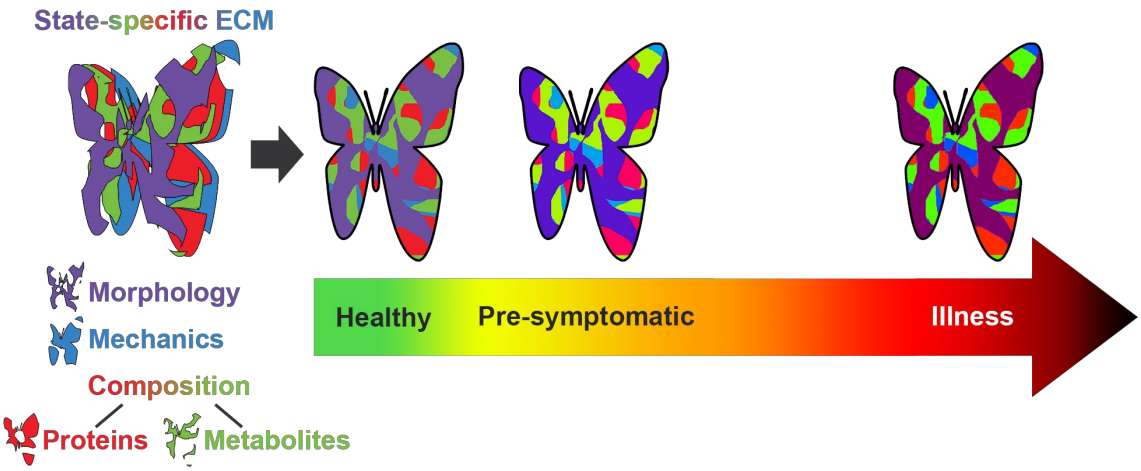
Fig. 5 Validation of identified biomarkers in human IBD

(A) Significantly different ECM proteins from MS analysis display similar trends in human dataset. Unpublished data from a proteomic data set of colon biopsies, Bennike *et al.*[79] were analyzed for changes in ECM protein abundance, comparing UC[79] 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 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 healthy donors vs. inflamed regions in UC patients. Controls do not demonstrate FBN1 staining (FBN1 negative), while a subset (n=5) 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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



Scheme 1 The ECM is an ultra-sensitive sensor of tissue states throughout disease progression, as revealed by multidimensional analysis of its biomaterial properties.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 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

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).

Samples were taken from both male and female subjects. The samples in each group for the RNA analysis are listed in **Table S8**. 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

Acute colonic inflammation was induced by administration of 1.25% DSS (MP Biomedicals LLC) in drinking water of C57BL/6 mice for seven days[56]. 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[60]. Mice were

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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[84]. 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). IL-10 mice lose weight only in early phases of colitis development, and the weight is regained as chronic colitis is established (**Fig. S1B**).

Mass cytometry and cell sorting

Cell isolation of colon samples was performed as previously described[85]. In brief, colons were harvested on the day of the experiment from healthy WT mice, along with those of day 4 and day 10 time points of the DSS model, were carefully separated from the surrounding fat and lymphatic vessels and cleaned from feces with cold PBS. Then, 2cm colon pieces were minced and immersed into RPMI-1640 medium (GIBCO, 21875034) containing 10% FBS, 2 mM

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

HEPES, 0.05mg DNase (Roche, 04716728001), and 0.5mg/mL of collagenase VIII (Sigma-Aldrich, C2139) and placed for 40 minutes at 250 rpm shaking at 37°C and mashed through a 70µm cell strainer.

Following, cells were stained according to a previously published protocol[86]. Individual mice cell suspensions were stained with Cell-ID Cisplatin 0.125 µM for viability and fixed using Maxpar® Fix I Buffer. Samples were then permeabilized using Maxpar® Barcode Perm Buffer and then barcoded using The Cell-ID™ 20-Plex Pd Barcoding Kit, allowing us to join samples for antigen staining. Antibodies used for staining are listed in **Table S1**. Before analyzing, the cell suspension was incubated with Cell-ID Interculator Iridium for 20 minutes.

Cells were analyzed with a cyTOF2® mass cytometer (Fluidigm). Results were normalized and debarcoded using Fluidigm cyTOF software[87]. CyTOF results gating and further analysis was done with FlowJo software (FlowJo, LLC) and cell populations were defined according to the markers listed in **Table S2**. tSNE analysis was done using the viSNE application in the Cytobank web platform. tSNE (t-Distributed Stochastic Neighbor Embedding) is a machine-learning algorithm used to cluster multivariate data into a 2D representation.

Decellularization of colonic tissue

Decellularization was performed as previously described[88]. 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.

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-

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 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 [89]. Then, 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[66] on unfixed 10μm 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 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 coverslips (diameter 15 mm) by means of a thin bi-adhesive tape. Samples were then attached to

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 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[90,91], recording the local deformation of the sample induced by the AFM probe. According to a recently published protocol[63], we used monolithic borosilicate glass probes consisting in micrometer-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[62].

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 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[63]. The values of the Young's Modulus (YM) were extracted by fitting the Hertz model to the experimental data[90,91]. 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 $\sigma_{med} = \sqrt{\pi/2}E_{med}\sigma_g^{10}/\sqrt{N}$ [92], 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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[93,94] (version 1.5.2.18) with the built-in Andromeda search engine[95] 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 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[96].

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1070 Computational analysis of MS data

Classification was done using a J48 decision tree algorithm[73] 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).

1075 Quantitative Real Time PCR on mouse and human colon tissues

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
1080 Biosystems inc.) on an ABI 7300 instrument (Applied Biosystems). Values are normalized to an 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 S7**. 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[97]. Mice samples were
1085 grouped according to general state: (i) healthy WT, (ii) pre-symptomatic (day 4 of DSS and 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
1090 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Immunohistochemical (IHC) stain for Fibrillin-1

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 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, 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 ± s.e.m) and statistical significance are reported and portrayed in 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[98]. 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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Statistical analysis of MS data

Bioinformatics analysis was performed with the Perseus program version 1.5.1.4[99]. 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[100]. 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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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.

Data 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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplemental Information for

Unraveling the state-specific nature of the native extracellular matrix via multidimensional characterization of its material properties

Elee Shimshoni[1], Ran Afik[1], Inna Solomonov[1], Idan Adir[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 S6
Tables S1 to S9

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

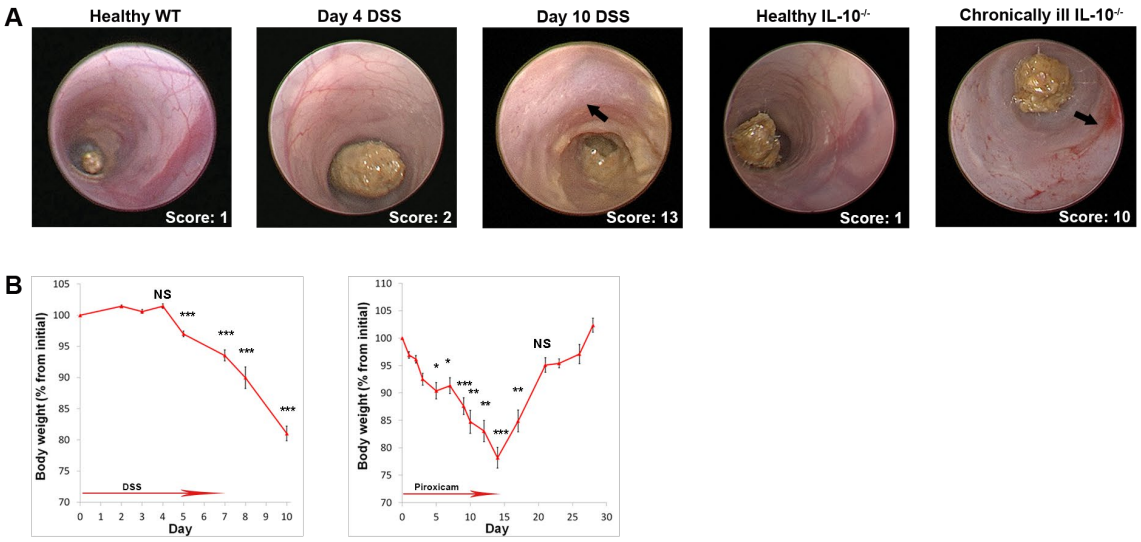
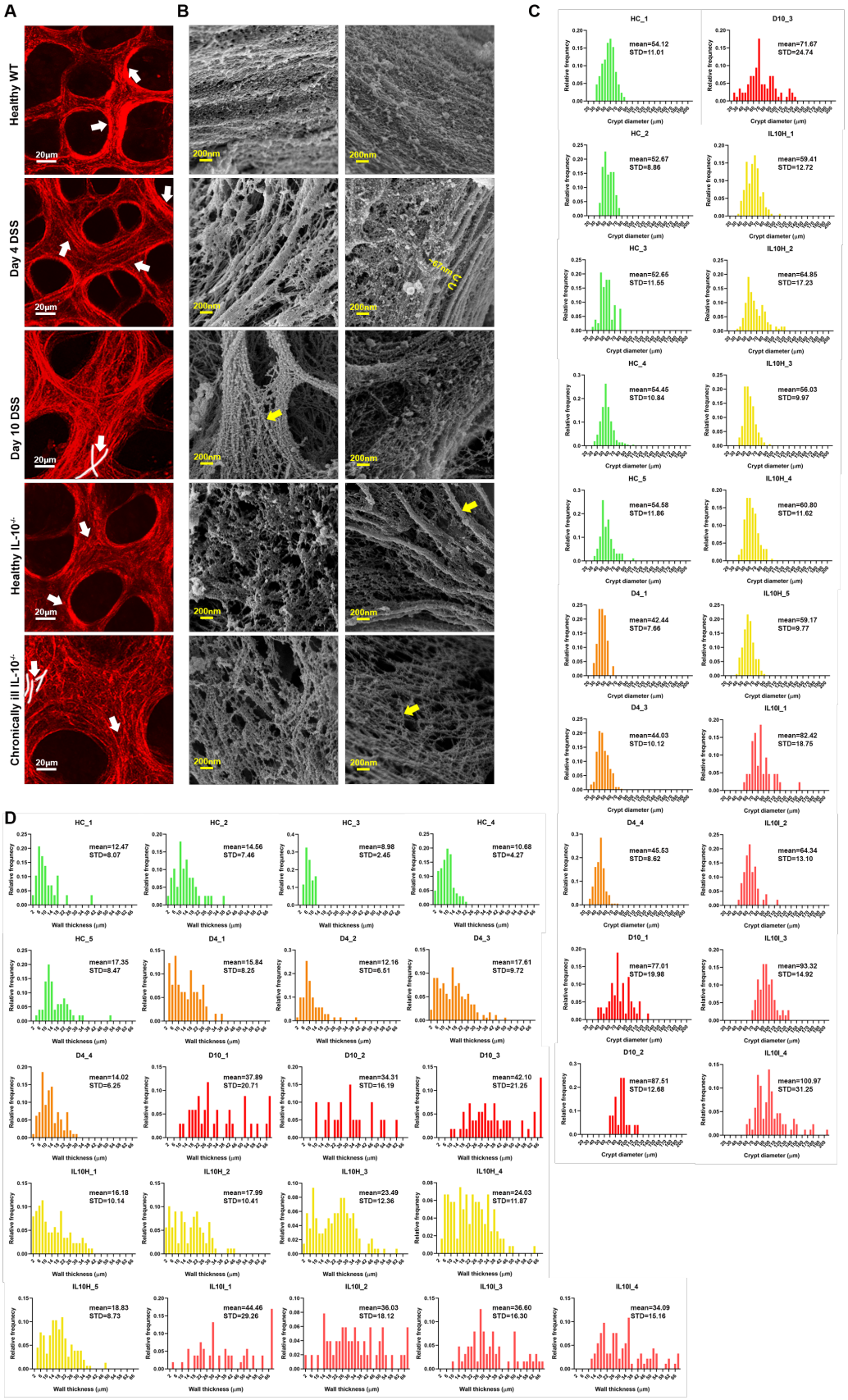


Fig. S1

(A) Endoscopic images. Live mice were imaged at the indicated states of the two models and inflammation scored as described in the Methods. 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) 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 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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

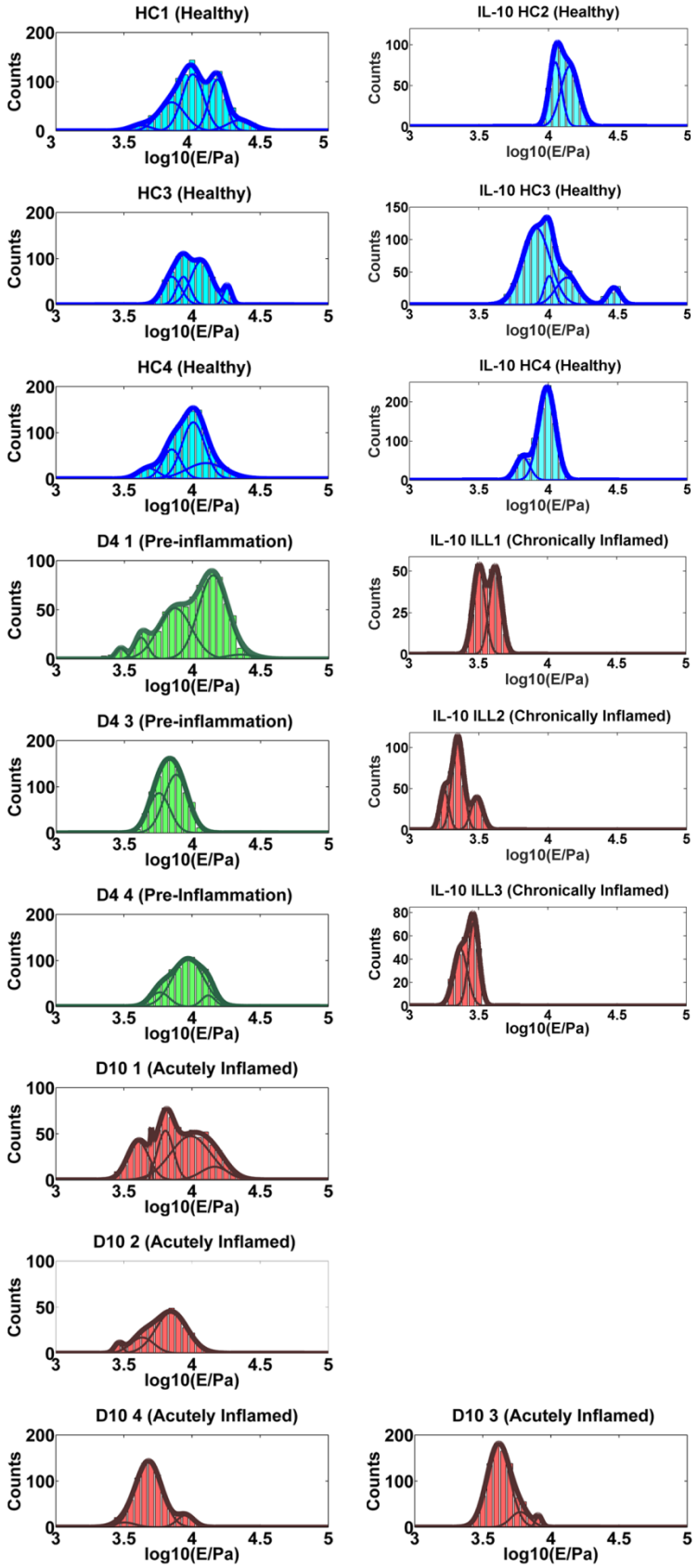


bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. S2

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 [101]. 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[61]. 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. (C+D) Histograms depicting the distributions of crypt diameters, in **C**, or wall thickness, in D, of each sample whose mean and STD can be found in the plots in **Fig. 2C**.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. S3

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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

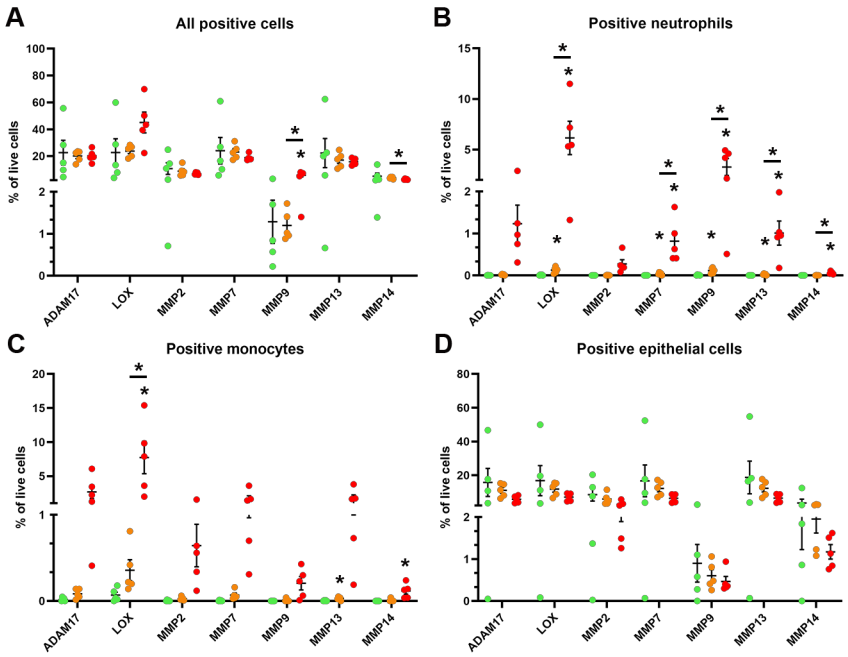


Fig. S4

Scatter dot plots based on the mass cytometry results depicting the changes in the relative frequencies (out of all live cells) of different cell populations positive for the indicated remodeling enzymes (A- all positive cells, B-neutrophils, C-monocytes and D-epithelial cells) at each state: healthy WT (green), day 4 of DSS (orange), day 10 of DSS (red). Number of animals: n=5 per state; Bonferroni correction for three comparisons: *P<0.0167, **P<0.00333.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

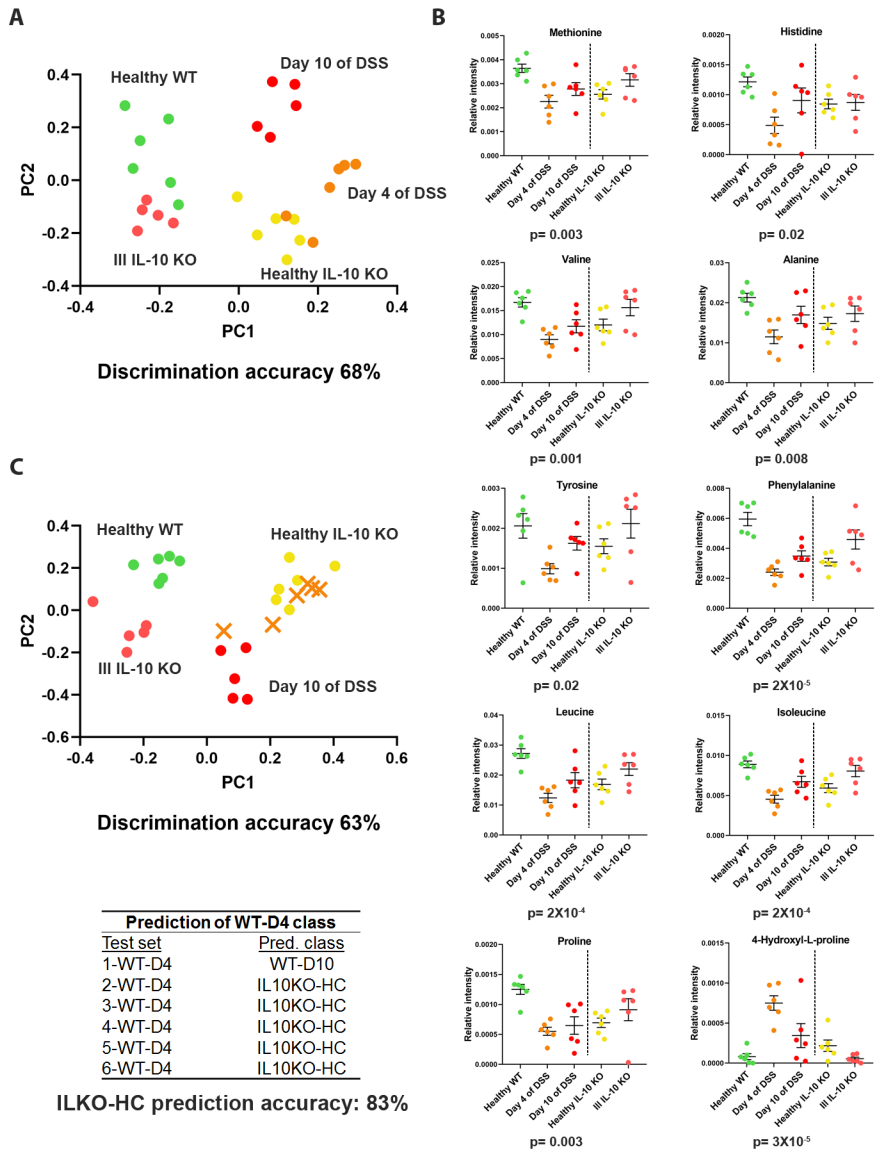


Fig. S5 Extracellular matrix proteolysis reveals its role as a metabolite reservoir that is modified under different inflammatory conditions

NMR-based untargeted metabolomic analysis on solutions with products of MMP-1 proteolysis of decellularized murine colon tissues in the five different states – healthy WT (green), day 4 of DSS (orange) healthy IL-10$^{-/-}$ (yellow), day 10 of DSS (red), ill IL-10$^{-/-}$ (pink). (A) PCA-CA analysis revealed that MMP-1 proteolysis resulted in differentiating metabolic profiles (which did not occur in the absence of the enzyme, see **Fig. S6**) of the degradation products (full list in

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table S9) across the five colitis states of the two models, with a 5-group discrimination accuracy of 68%. (B) Scatter dot plots of metabolites displaying statistically significant different levels in the five states with the indicated one-way ANOVA p values. Pre-symptomatic states display reduced extraction of amino acids upon proteolysis compared to both healthy WT and the two ill states. This indicates that the pre-symptomatic ECM is a poor reservoir of these types of metabolites. However, most interestingly, DSS day 4 matrix 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. DSS day 4 samples are predicted as the other pre-symptomatic state – healthy IL-10$^{-/-}$. In the score plots each symbol represents a different sample and each color a different group of mice.

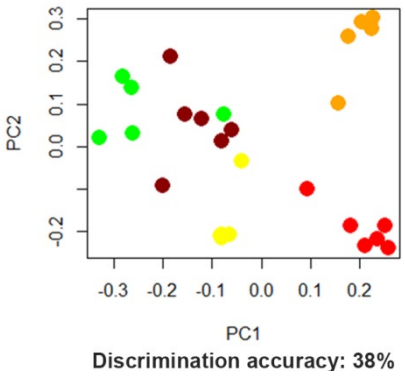


Fig. S6

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 without MMP addition (38% discrimination accuracy).

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Name	Product	Provider	Isotope conjugation
Anti-Human/ Mouse/ Rat Ki-67	3168007	FLUIDIGM	168Er - FLUIDIGM
Anti-Human/Mouse CD45R/B220	3160012	FLUIDIGM	160Gd – FLUIDIGM
Anti-Mouse CD11c	3142003	FLUIDIGM	142Nd – FLUIDIGM
Anti-Mouse NK1.1	3170002	FLUIDIGM	170Er – FLUIDIGM
Anti-Mouse CD4	3145002	FLUIDIGM	145Nd – FLUIDIGM
Anti-Mouse CD8a	3146003	FLUIDIGM	146Nd – FLUIDIGM
Anti-Mouse TER-119	3154005	FLUIDIGM	154Sm – FLUIDIGM
Anti-Mouse Ly-6C	3162014	FLUIDIGM	162Dy – FLUIDIGM
Anti-Mouse Ly-6G	3141008B	FLUIDIGM	141Pr – FLUIDIGM
Anti-Mouse CD64	3151012	FLUIDIGM	151Eu – FLUIDIGM
Anti- Mouse F4/80 (BM8)	3159009	FLUIDIGM	159Tb – FLUIDIGM
Anti-Mouse CD31/PECAM-1 (390)	3165013	FLUIDIGM	165Ho – FLUIDIGM
Anti- Mouse CD326 [EpCAM] (G8.8)	3166014	FLUIDIGM	166Er – FLUIDIGM
Anti-Mouse CD45	3089005	FLUIDIGM	89Y – FLUIDIGM
Anti-mouse CD3e (maxpar ready)	BLG-100345	Biolegend	153Eu – FLUIDIGM
Anti-Mouse CX3CR1 (SA011F11)	3164023	FLUIDIGM	164Dy – FLUIDIGM
Anti-Mouse CD206	3169021B	FLUIDIGM	169Tm – FLUIDIGM
Anti-Mouse I-A/I-E (M5/114.15.2)	3209006B	FLUIDIGM	209Bi – FLUIDIGM
Anti-Mouse CD11b	3143015	FLUIDIGM	143Nd – FLUIDIGM
Anti-MMP9	ab38898	Abcam	149Sm – Home made
Anti-MMP14	ab51074	Abcam	156Gd – Home made
Anti-TIMP2	ab38973	Abcam	172Yb – Home made
Anti-MMP2	ab92536	Abcam	176Yb – Home made
Anti-ADAM-17	ab57484	Abcam	174Yb – Home made
Anti-MMP7	ab5706	Abcam	148Nd – Home made
Anti-CD34	14-0341-82	Invitrogen	152Sm – Home made
Anti-LOX	ab31238	Abcam	175Lu – Home made
Anti-αSMA	ab32575	Abcam	173Yb – Home made
Anti-MMP-13	ab39012	Abcam	150Nd – Home made

Table S1

List of antibodies used in the mass-cytometry analysis.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ID	Markers
Epithelial cells	CD326+ ; CD45-
Early endothelial progenitors	CD326- ; CD45- ; CD34+ ; CD31-
Late endothelial progenitors	CD326- ; CD45- ; CD34+ ; CD31+
Adult Endothelial cells	CD326- ; CD45- ; CD34- ; CD31+
Cytotoxic T lymphocytes	CD326- ; CD45+ ; CD11b- ; CD3e+ ; CD8a+ ; CD4-
T helper lymphocytes	CD326- ; CD45+ ; CD11b- ; CD3e+ ; CD8a- ; CD4+
B cells	CD326- ; CD45+ ; CD11b- ; CD3e- ; CD45R+
Dendritic cells	CD326- ; CD45+ ; CD11c+ ; CD3e- ; CD45R- ; MHC class II +
Neutrophils	CD326- ; CD45+ ; CD11b+ ; CD3e- ; CD45R- ; Ly-6G+
Macrophages	CD326- ; CD45+ ; CD11b+ ; CD3e- ; CD45R- ; Ly-6G- ; F4/80+
Circulating (pro inflammatory) monocytes	CD326- ; CD45+ ; CD11b- ; CD3e- ; CD45R- ; Ly-6G- ; F4/80- ; Ly-6Chi

Table S2

Gating and cell population definition in the mass cytometry analysis.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 S3

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. S3**, which are plotted in **Fig.2E-F**. 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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 S4

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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 S5

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.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 S6

List of proteins in the three different clusters from **Fig. 4G** according to their abundance pattern across samples.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 S7

Primer sequences used for quantitative real-time PCR analysis.

1315

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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

Table S8

List of human samples in each group of mRNA analysis displayed in **Fig. 5C**.

bioRxiv preprint doi: https://doi.org/10.1101/665653; this version posted July 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 S9

List of metabolites released from proteolytically cleaved ECMs.