Sex-dimorphic expression of extracellular matrix genes in mouse bone marrow neutrophils

Cassandra J. McGill[1], Collin Y. Ewald[2, §] and Bérénice A. Benayoun[1,3,4,5,6,§]

[1]Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA.

[2] Laboratory of Extracellular Matrix Regeneration, Institute of Translational Medicine, Department of Health Sciences and Technology, Swiss Federal Institute of Technology (ETH Zürich), Schwerzenbach, Switzerland.

[3] Molecular and Computational Biology Department, USC Dornsife College of Letters, Arts and Sciences, Los Angeles, CA 90089, USA.

[4] Biochemistry and Molecular Medicine Department, USC Keck School of Medicine, Los Angeles, CA 90089, USA.

[5] USC Norris Comprehensive Cancer Center, Epigenetics and Gene Regulation, Los Angeles, CA 90089, USA.

[6] USC Stem Cell Initiative, Los Angeles, CA 90089, USA.

[§]co-corresponding authors

Abstract

The mammalian innate immune system is sex-dimorphic. Neutrophils are the most abundant leukocyte in humans and represent innate immunity's first line of defense. We previously found that primary mouse bone marrow neutrophils show widespread sex-dimorphism throughout life, including at the transcriptional level. Extracellular matrix [ECM]-related terms were observed among the top sex-dimorphic genes. Since the ECM is emerging as an important regulator of innate immune responses, we sought to further investigate the transcriptomic profile of primary mouse bone marrow neutrophils at both the bulk and single-cell level to understand how biological sex may influence ECM component expression in neutrophils throughout life. Here, using curated gene lists of ECM components and unbiased weighted gene co-expression network analysis [WGCNA], we find that multiple ECM-related gene sets show widespread female-bias in expression in primary mouse neutrophils. Since many immune-related diseases (*e.g.,* rheumatoid arthritis) are more prevalent in females, our work may provide insights into the pathogenesis of sex-dimorphic inflammatory diseases.

Introduction

Accumulating evidence shows that, even outside of reproductive organs, mammalian biology is very sex-dimorphic [1]. Indeed, the mammalian immune system shows broad differences between males vs. females, including aspects of innate and adaptive immune responses [2-4]. Generally, a more robust immune response is observed in females, which may underlie a higher auto-immunity risk, whereas an increased susceptibility to infection and a worse response to sepsis is found in males [2].

Neutrophils are the most abundant cell type in human blood, representing 50-70% of leukocytes in humans throughout life [5]. They are a type of “granulocytes”, produced in the bone marrow and released into circulation for immune surveillance, representing a first line-of-defense against infections [6, 7]. Neutrophils perform key immune functions, including the production/release of antimicrobial granules and of “Neutrophil Extracellular Traps” (NETs) [6, 8], composed of extruded chromatin, proteases, and antimicrobial peptides. While neutrophils defend against infections, their aberrant activation aggravates pathological inflammation [6, 7, 9]. They can also drive inflammatory disease, leading to tissue damage [8, 10].

Recent work by our group revealed widespread sex-dimorphism in the transcriptome of primary murine bone marrow neutrophils throughout life (4 to 20 months of age) [11]. Consistently, transcriptomic data by the ImmGen Consortium showed clear differences between the neutrophils of young prepubertal female *vs.* male mice (6 weeks of age) [4]. Importantly, we also recently reported that sex-differences in primary bone marrow neutrophils are found at every stage of underlying heterogeneity using single-cell RNA-seq [12]. In our previous analyses of primary neutrophil transcriptomes, Gene Ontology [GO] terms related to collagen synthesis and extracellular matrix (ECM) biology were observed in the top listed GO terms with significant sex-dimorphic gene expression at the bulk RNA-seq level [11]. Furthermore, beyond the transcriptional level, accumulating evidence supports the idea that the functional landscape of neutrophils differs between sexes throughout life, with male-bias in the production of primary granule protein neutrophil elastase (ELANE) [11], a serine protease that can break down elastin and collagens [13], potentially remodels the ECM.

Intriguingly, ECM components are emerging as potent regulatory mediators of innate immune responses [14]. The importance of the extracellular matrix in neutrophil recruitment and function is becoming more evident based on recent studies [15]. Indeed, recent work suggests that neutrophils can transport ECM components to wound sites to help promote healing [16]. Although the production of ECM components by neutrophils themselves has not been explored in depth, emerging evidence suggests that the synthesis of “emergency” ECM containing

fibronectins by neutrophils plays an important role in promoting fracture healing [17]. However, sex-differences in the regulation of ECM biology by neutrophils remain poorly understood.

To follow-up on this tantalizing observation of sex-dimorphism in ECM gene expression, we reanalyze transcriptomic data from primary mouse bone marrow neutrophils to help understand how biological sex may influence ECM component gene expression. Using expert-curated ECM gene sets (*i.e.,* matrisome), we confirm broad female-bias in the expression of genes related to ECM components and remodeling in young and old female neutrophils in our bulk RNA-seq dataset. To gain a more granular view and determine any underlying heterogeneity of ECM expression of young mice's neutrophils, we recently generated a single cell RNA-seq data and find again a female-specific enrichment for ECM. In addition, using unbiased weighted gene co-expression network analysis (WGCNA), we identify significant sex-biased gene expression modules in primary neutrophils. Importantly, the top female-biased WGCNA module in neutrophils was significantly enriched for ECM-related genes. Thus, our analyses suggest that female neutrophils may uniquely contribute to tissue repair and ECM remodeling, in addition to their core function in the immune response.

Results

Gene set enrichment analysis confirms female-biased expression of ECM-related genes expression in primary mouse neutrophils

In our previous study using a multi-omic approach to assess changes in neutrophils based on age and sex, we observed a significant female bias (regardless of age) for expression of genes related to GO terms “Collagen chain trimerization”, “Assembly of collagen fibrils and other multimeric structures”, and “Non-integrin membrane-ECM interactions” [11]. Because of the emerging role of neutrophils in ECM-remodeling [15], we decided to investigate further the potential differential expression of extracellular matrix (ECM) related processes, as a function of organismal aging or biological sex.

To determine whether ECM-related genes are differentially regulated as a function of organismal age or biological sex, we took advantage of 2 datasets that we previously generated: (i) a bulk transcriptomic dataset of primary bone marrow neutrophils derived from young and old, female and male C57BL/6Nia mice [11], and (ii) a single cell transcriptomic dataset of primary bone marrow neutrophils derived from young female and male C57BL/6J mice [12] (**Figure 1A**). In addition, to maximize our understanding of the differential expression of ECM-related genes in neutrophils, we leveraged expert-curated gene sets related to ECM biology (*i.e.,* matrisome; **Supplementary Table S1**). The matrisome encompasses all proteins that form the extracellular matrix (*i.e.*, collagens, ECM glycoproteins, proteoglycans), also referred to as the core matrisome, encompassing 274 genes in mice [18]. Furthermore, the matrisome also includes proteins that either remodel the ECM (*i.e.,* matrix metalloproteases or ECM regulators) or localize into the ECM (*i.e.,* secreted factors, ECM-affiliated genes), collectively referred to as matrisome-associated genes, which are 836 genes in mice [18]. Thus, examining the changes in the composition of these 1,110 matrisome genes might help us reveal some new underlying biology.

First, we leveraged these gene sets to perform Gene Set Enrichment Analysis (GSEA) on bulk transcriptomic neutrophil data to detect potential changes as a function of aging or sex (**Figure 1B; Supplementary Table S2A-B**). We observed no significant changes in the expression of ECM-related genes based on organismal age, suggesting that ECM biology is not broadly disrupted in aging mouse neutrophils (**Supplementary Table S2A**). By contrast, we did find robust female bias in the expression of most ECM-related gene sets in neutrophils (**Figure 1B**; **Supplementary Table S2B**). In addition to a female-biased enrichment in collagen-encoding genes (**Figure 1B**, FDR = 4.26×10^{-9}), we also observed significant female-biased expression for genes related to the core matrisome, ECM affiliated genes, ECM glycoproteins,

matrisome-associated genes, and secreted factors, but not for proteoglycan-related genes (**Figure 1B**; **Supplementary Table S2B**). Using Fisher's exact test for the significance of overlap, we confirmed that genes from ECM-related gene sets tended to be overrepresented among genes with significant female-biased neutrophil expression by bulk RNA-sequencing (DESeq2 FDR < 5%; **Supplementary Figure S1A; Supplementary Table S2C**). Interestingly, genes expressed at significantly higher levels in female neutrophils show 1.7- to 7.9-fold enrichment for ECM-related genes compared to chance (**Supplementary Table S2C**).

We were also curious whether the female ECM-related gene expression bias would also be observed at the single-cell level. For this purpose, we took advantage of a recent single-cell RNA-seq dataset from female *vs*. male mouse bone marrow neutrophils [12]. To minimize the impact of "dropouts" in single-cell RNA-seq datasets [19], we used a gene set scoring approach, UCell [20], to determine the overall expression of the ECM-related gene set in each individual cell. This approach computes a single numeric value per cell based on the expression of genes in a gene set of interest without the need for expression thresholding, thus increasing statistical power [20]. We show the summarized the UCell scoring values for each ECM-related gene set per biological sample, and the significance of the difference in score distributions in female *vs.* male neutrophils (**Figure 1C**). Importantly, this analysis revealed a robust female-bias expression for most of the ECM-related gene sets (**Figure 1C**). Interestingly, based on neutrophil subset definitions from [21], ECM-related genes may be highest in more mature neutrophils, suggesting that this phenotype is acquired during the maturation process (**Supplementary Figure S1B**).

Interestingly, although collagens were not robustly detected in single-cell RNA-seq due to relatively low expression and sensitivity of the assay (<10% of cells showed detectable collagen transcript levels in single-cell RNA-seq regardless of sex; **Figure 1C**), collagen genes showed robust female-bias in their expression at the bulk transcriptome level (**Figure 1D, E**).

Although we could not find a publicly available mouse neutrophil quantitative proteomics dataset with female and male animals, we identified a high-quality serum proteomics dataset derived from young C57BL/6NHsd female and male mice [22]. Since neutrophils represent a large proportion of circulating blood leukocytes [5], we reasoned that differences in the production of ECM-related proteins by neutrophils should be at least partially reflected in the serum proteome. Thus, we identified significantly sex-biased serum proteins using Limma [23], with respectively 73 female-biased and 42 male-biased serum proteins at FDR < 5% (**Supplementary Figure S1C-D**; **Supplementary Table S2D**). Consistent with our

transcriptomics observations, ECM-related gene sets tend to be enriched among serum female-biased proteins (Limma FDR < 5%; **Supplementary Figure S1E; Supplementary Table S2E**).

Overall, our analysis determined that genes that both encode structural components of the ECM and remodel the ECM tend to be expressed at higher levels in female neutrophils throughout life (and found at higher protein levels in the serum), suggesting that the emerging interactions between neutrophil activity and ECM biology may be sex-dimorphic throughout life.

Weighted gene co-expression network analysis reveals that the top female-biased module in neutrophil transcriptomes is enriched for collagens

We next asked which gene modules were coregulated together with ECM-related genes in a female-biased fashion by leveraging weighted gene co-expression network analysis (WGCNA) (**Supplementary Figure S2A**). Since WCGNA is optimized to be used on bulk datasets, we focused this analysis on the bulk transcriptomic neutrophil data. After network construction, we identified 13 significant transcriptional modules (**Supplementary Table S3**). We next used the WGCNA module-trait relationship feature to identify which modules were significantly associated with organismal age and/or sex (**Figure 2A**).

Most notably, the most significant sex-associated module, the Salmon module, which contains 318 genes, showed clear female-bias regardless of age (**Figure 2B, Supplementary Figure S2B**). To note, there was no significant enrichment for X-linked genes in the Salmon module (p = 0.9153 in Fisher's exact test), although some X-linked genes were found in this module (*i.e., Xist, Tsix, Nyx, Kdm6a, Kdm5c, Zxdb)*. Manual inspection also revealed that this module contains many collagen genes (**Supplementary Table S3**). By contrast, the module with the second most significant sex-association, the Magenta module, which contains 544 genes, showed male-bias, as well as decreases with aging (**Figure 2B**). To unbiasedly determine which functions were associated with the Salmon and Magenta module, we used hypergeometric enrichment with Gene Ontology and Reactome categories (**Figure 2C, D**; **Supplementary Table S3**). Consistent with the presence of collagen-related genes, the Salmon module was strongly associated with collagen- and ECM-related terms (**Figure 2C, D**; **Supplementary Table S3**). By contrast, the Magenta module was strongly associated with chromatin and cell cycle-related terms (**Supplementary Table S3**), consistent with our previous observations that these terms were male-biased and regulated with aging [11]. Together, our analysis of the topmost sex-associated WGCNA modules revealed that ECM gene expression is a sex-dimorphic transcriptional feature of neutrophils.

187

Discussion

While neutrophils, upon activation, are known to secrete elastase and MMPs to remodel ECM [13, 15], activate fibroblasts [24, 25], or even transport ECM to wound sites [16], little is known about the intrinsic expression of ECM genes of neutrophils in their bone marrow niche. Here, we show that the most significant gene cluster correlating with biological sex (Salmon cluster) is enriched with a distinct set of ECM genes that are biased to be expressed at higher levels in female neutrophils. With our analysis, we were not able to detect an impact of aging on the regulation of ECM-related gene expression (regardless of sex). Nevertheless, it is possible that changes related to aging as a function of sex may emerge in datasets with larger numbers of animals, which should be investigated in future studies.

We identified potential gene sets in the Salmon cluster that might prime female neutrophils for improved wound healing (**Supplementary Table S3**). This gene network consists of collagens interacting with platelet-derived growth factor binding (**Figure 2**), known to initiate neutrophil activation and inflammatory responses at sites of injury [26]. Interestingly, integrin a5 (Itga5) was uniquely expressed in the Salmon cluster. Neutrophils that express α5β1 integrin bind fibronectin [27]. Fibronectin is important for wound healing and is required to be placed first in order to add collagen to build a new functional ECM [28]. In human bone fracture hematomas, neutrophils localize to fibronectin shortly after injury, and hematoma-localized neutrophils are uniquely stained for cellular fibronectin [17]. This suggests that activated neutrophils in the hematoma may produce a first “emergency ECM” [17], consistent with our observation of fibrillar collagen gene expression of neutrophils. Furthermore, upon adhering to fibronectin, neutrophils secrete branched-chain (valine, isoleucine, leucine), aromatic (tyrosine, phenylalanine), and positively charged free amino acids (arginine, ornithine, lysine, hydroxylysine, histidine) [17]. Although the reason for this is unclear, the release of, for instance, hydroxylysine, crucial for Lysol oxidase-mediated collagen crosslinking, is either increased by stimulating neutrophils with insulin or reduced by blocking the PI3K/Akt pathway [17]. Based on this and the composition of the female-enriched Salmon module, we speculate on a connection among this metabolic axis via serum/glucocorticoid regulated kinase 1 (Sgk1), integrin a5, and fibrillar collagens (Col1a1/2, Col2a1), integrated stress response activating transcription factor 4 (Atf4), also implicated in amino acid metabolism for collagen biosynthesis [29], and potentially modulating immune responses via Pdgfa and interleukin 6 receptor alpha (Il6ra) and further collagen synthesis fibroblast growth factor receptor 1 (Fgfr1) (**Supplementary Table S3**). This hypothesis might be interesting to test in the future and might provide a mechanistic explanation for why female mice show enhanced wound-healing compared to males [30, 31].

Understanding neutrophil biology has broad implications for therapeutic interventions. Additionally, investigating sex-dimorphic mechanisms affecting neutrophil responses is vital for developing interventions beneficial to both males and females. For example, there is a clear female bias when it comes to rheumatoid arthritis, whereby 75% of patients are women [32]. Inducing arthritis in mice by anti-type II collagen antibodies and lipopolysaccharide, leads to massive neutrophil infiltration into the joint space [33]. Fascinatingly, depleting neutrophils either by the onset of arthritis or during arthritis completely ameliorates the disease [33], suggesting that neutrophils not only initiate but are also involved in maintaining arthritis disease progression. Thus, investigating female-biased processes in neutrophils, including the contribution of the ECM, might be an important next step in managing rheumatoid arthritis.

In summary, we examined the sex-specific gene expression differences of neutrophils. We established that a major difference between male and female neutrophils is the expression of ECM genes. We identified a gene network that might prime ECM gene expression in female neutrophils to potentially accelerate wound healing. Dissecting sex-specific and intrinsic gene expression of neutrophils may significantly impact our understanding to tailor sex-specific therapeutic interventions for wound healing, immune responses, and arthritis.

Methods

Data acquisition and preprocessing

We obtained published processed bulk RNA-seq differential expression results from young (4 months) and aged (20 months), male and female primary bone marrow neutrophils derived from C57BL/6JNia mice (PRJNA630663; **Figure 1A** upper) [11]. For all analyses below, we used the SVA-normalized count table and DEseq2 analyses results generated in that study, for which the processed files were available online from https://github.com/BenayounLaboratory/Neutrophil_Omics_2020 [11].

We also obtained a processed and annotated Seurat object containing data from a 10xGenomics single-cell RNA-seq dataset of young (3 months) male and female primary bone marrow neutrophils derived from C57BL/6J mice (PRJNA796634; **Figure 1A** lower) [12]. This file was obtained from Figshare (https://doi.org/10.6084/m9.figshare.19623978) [12]. R package 'Seurat' 4.2.0 was used to load and interact with the annotated single-cell RNA-seq expression dataset.

Finally, we obtained a processed serum proteomics dataset from C57BL/6NHsd female and male mice (4-5 months of age) [22]. The processed proteomics expression file was obtained from Table S2 from the publisher's site. For our analysis, we only considered label-free quantitation [LFQ] intensity data derived from wild-type animals. We performed Multidimensional Scaling (MDS) analysis on protein expression levels in serum using a distance metric between samples based on the Spearman's rank correlation value (1-Rho), then provided to the core R command 'cmdscale'. Limma 3.50.3 [23] was used to perform differential expression analysis, with sex as the variable of interest and ascorbate treatment as a modeling covariate. Proteins with Limma False Discovery Rate [FDR] < 0.05 were considered significantly sex-biased and are reported in **Supplementary Table S2D**.

Curated lists of extracellular matrix-related terms

We obtained a manually curated list of gene sets relevant to extracellular matrix [ECM] biology [18]. The curated gene sets are also provided in **Supplementary Table S1.**

Gene Set Enrichment Analysis (GSEA) of extracellular matrix-related terms

We used the GSEA paradigm to determine whether ECM-related gene sets were differentially regulated as a function of sex or age in bulk RNA-seq of primary neutrophils[34]. For this purpose, we used R package 'phenotest' 1.42.0 in R version 4.1.2. The DEseq2 t-statistic for each gene calculated for age- or sex-expression was used to create the ranked

gene list for functional enrichment analysis for both sex and aging [11]. The table output of this analysis is reported in **Supplementary Table S2A-B** and graphically represented in **Figure 1B**.

Overrepresentation analysis using Fisher's exact test

To determine the significance of the overlap between significantly female-biased genes in neutrophils by RNA-seq (DEseq2 FDR < 0.05) or serum proteomics (limma FDR < 0.05) and ECM-related gene lists, we used Fisher's exact test. The background used was all genes detected by RNA-seq or proteins detected by proteomics, respectively. P-value correction for multiple hypothesis testing was performed using the Benjamini-Hochberg method. The table output of this analysis is reported in **Supplementary Table S2C,E** and graphically represented in **Figure S1A,E**.

To determine the significance of the overlap between X-linked genes and genes from the Salmon module, we obtained a list of X-encoded genes from ENSEMBL BioMart version 108 (database access 2023/02/02), restricting output based on gene position on chromosome X. This list was further restricted to genes expressed in our bulk neutrophil RNA-seq according to our previous study [11]. The Fisher's exact test was used to determine the significance of the overlap between genes of the Salmon module and expressed X-linked genes, using the list of expressed genes from the bulk RNA-seq as the universe.

Ucell scoring analysis of extracellular matrix-related terms

To determine whether ECM-related terms were differentially regulated as a function of neutrophil biological sex at the single-cell level, we leveraged the UCell robust single-cell gene signature scoring metric implemented through R package 'UCell' 1.3.1 [20]. Cell-wise UCell scores were computed for each ECM-related gene list and are graphically reported as aggregates by sample of origin in **Figure 1C**. Note that Collagen-related genes were not robustly detected at the single cell level, likely a limitation of the platform, and have an extremely low percentage of cells with detected expression (**Figure 1C**). For analysis of statistical significance, a non-parametric Wilcoxon rank sum test was used to compare the distribution of UCell scores between female and male neutrophils and is reported for each gene set in **Figure 1C**.

Weighted gene co-expression network analysis [WGCNA]

To investigate what other genes are being co-expressed with the collagens and other ECM-related genes, we used the WGCNA paradigm [35, 36]. We used R package 'WGCNA'

1.71 [36] and the above-mentioned normalized SVA count table as the inputs. The data was preprocessed by removing any genes with inconsistent expression, as recommended by the user manual. A power of 10 was chosen to create the scale-free topology based on the power and scale independence equilibration. The produced topological overlap matrix [TOM] reveals the mean network connectivity of each gene, and genes with similar expression profiles were then classified into different modules. We set the minimum module size to 100. The analyses yield 13 significant modules. We then performed module-trait analysis using age and sex as the traits to correlate to the modules.

Functional enrichment analysis of WGCNA gene modules

We performed functional enrichment analysis for genes associated with the top sex-biased module, the 'Salmon' module ($p = 2x10^{-10}$; **Figure 2A-B**), to unbiasedly determine the top functions differentially regulated in primary neutrophils as a function of sex. We used a hypergeometric test to determine overrepresented Gene Ontology [GO] terms and REACTOME Genesets in the Salmon module compared to expressed genes in the bulk neutrophil transcriptome dataset. For this purpose, we took advantage of R package 'ClusterProfiler' 4.2.2 [37], with GO annotations derived from Bioconductor annotation package 'org.Mm.eg.db' 3.14.0, and with Reactome annotations derived from the 'Reactome-PA' 1.38.0 package [38]. Both dotplot and gene concept network results were derived from ClusterProfiler plotting functions.

Acknowledgements

Some panels created with BioRender.com. This work was supported by NIA T32 AG052374 predoctoral fellowship to C.J.M., by the Swiss National Science Foundation Funding from the SNF P3 Project 190072 to CYE, NIA R01 AG076433 to B.A.B. and Pew Biomedical Scholar award #00034120 to B.A.B.

Author Contributions

C.Y.E. and B.A.B designed the study. C.J.M. and B.A.B. performed data analyses. All authors contributed to the writing and editing of the manuscript.

Data Availability

The bulk RNA-seq data was previously described in Lu et al., 2021 [11], and sequencing data is accessible through BioProject PRJNA630663. The processed normalized count table and DEseq2 result tables were obtained from Github (https://github.com/BenayounLaboratory/Neutrophil_Omics_2020). The single-cell RNA-seq data was previously described in Kim et al., 2022 [12], and raw sequencing data is accessible through BioProject PRJNA796634. The processed annotated Seurat file was obtained from Figshare (https://doi.org/10.6084/m9.figshare.19623978). The serum proteomics dataset was previously described in Aumailley et al, 2021 [22], and the processed protein expression matrix was obtained from the online Table S2 for the article (https://pubs.acs.org/doi/suppl/10.1021/acs.jproteome.1c00542/suppl_file/pr1c00542_si_006.xlsx).

Code Availability

All the analytical code for this study is available on the Benayoun lab github (https://github.com/BenayounLaboratory/Neutrophil_ECM).

Competing Interests

The authors declare no competing interests.

References

[1] N.K. Sampathkumar, J.I. Bravo, Y. Chen, P.S. Danthi, E.K. Donahue, R.W. Lai, R. Lu, L.T. Randall, N. Vinson, B.A. Benayoun, Widespread sex dimorphism in aging and age-related diseases, Hum Genet 139(3) (2020) 333-356.

[2] S.L. Klein, K.L. Flanagan, Sex differences in immune responses, Nat Rev Immunol 16(10) (2016) 626-38.

[3] E.J. Marquez, C.H. Chung, R. Marches, R.J. Rossi, D. Nehar-Belaid, A. Eroglu, D.J. Mellert, G.A. Kuchel, J. Banchereau, D. Ucar, Sexual-dimorphism in human immune system aging, Nat Commun 11(1) (2020) 751.

[4] S.T. Gal-Oz, B. Maier, H. Yoshida, K. Seddu, N. Elbaz, C. Czysz, O. Zuk, B.E. Stranger, H. Ner-Gaon, T. Shay, ImmGen report: sexual dimorphism in the immune system transcriptome, Nat Commun 10(1) (2019) 4295.

[5] E.H. Nah, S. Kim, S. Cho, H.I. Cho, Complete Blood Count Reference Intervals and Patterns of Changes Across Pediatric, Adult, and Geriatric Ages in Korea, Ann Lab Med 38(6) (2018) 503-511.

[6] B. Shah, N. Burg, M.H. Pillinger, Chapter 11 - Neutrophils, in: G.S. Firestein, R.C. Budd, S.E. Gabriel, I.B. McInnes, J.R. O'Dell (Eds.), Kelley and Firestein's Textbook of Rheumatology (Tenth Edition), Elsevier2017, pp. 169-188.e3.

[7] R.C. Furze, S.M. Rankin, Neutrophil mobilization and clearance in the bone marrow, Immunology 125(3) (2008) 281-288.

[8] G. Sollberger, D.O. Tilley, A. Zychlinsky, Neutrophil Extracellular Traps: The Biology of Chromatin Externalization, Developmental cell 44(5) (2018) 542-553.

[9] O. Soehnlein, S. Steffens, A. Hidalgo, C. Weber, Neutrophils as protagonists and targets in chronic inflammation, Nat Rev Immunol 17(4) (2017) 248-261.

[10] V. Papayannopoulos, Neutrophil extracellular traps in immunity and disease, Nat Rev Immunol 18(2) (2018) 134-147.

[11] R.J. Lu, S. Taylor, K. Contrepois, M. Kim, J.I. Bravo, M. Ellenberger, N.K. Sampathkumar, B.A. Benayoun, Multi-omic profiling of primary mouse neutrophils predicts a pattern of sex and age-related functional regulation, Nat Aging 1(8) (2021) 715-733.

[12] M. Kim, R.J. Lu, B.A. Benayoun, Single-cell RNA-seq of primary bone marrow neutrophils from female and male adult mice, Sci Data 9(1) (2022) 442.

[13] S.E. McGowan, P.J. Stone, G.L. Snider, C. Franzblau, Alveolar macrophage modulation of proteolysis by neutrophil elastase in extracellular matrix, Am Rev Respir Dis 130(5) (1984) 734-9.

[14] A. Garcia-Garcia, I. Martin, Extracellular Matrices to Modulate the Innate Immune Response and Enhance Bone Healing, Front Immunol 10 (2019) 2256.

[15] Y. Zhu, Y. Huang, Q. Ji, S. Fu, J. Gu, N. Tai, X. Wang, Interplay between Extracellular Matrix and Neutrophils in Diseases, J Immunol Res 2021 (2021) 8243378.

[16] A. Fischer, J. Wannemacher, S. Christ, T. Koopmans, S. Kadri, J. Zhao, M. Gouda, H. Ye, M. Muck-Hausl, P.W. Krenn, H.G. Machens, R. Fassler, P.A. Neumann, S.M. Hauck, Y. Rinkevich, Neutrophils direct preexisting matrix to initiate repair in damaged tissues, Nat Immunol 23(4) (2022) 518-531.

[17] O.W. Bastian, L. Koenderman, J. Alblas, L.P. Leenen, T.J. Blokhuis, Neutrophils contribute to fracture healing by synthesizing fibronectin+ extracellular matrix rapidly after injury, Clin Immunol 164 (2016) 78-84.

[18] A. Naba, K.R. Clauser, H. Ding, C.A. Whittaker, S.A. Carr, R.O. Hynes, The extracellular matrix: Tools and insights for the "omics" era, Matrix Biol 49 (2016) 10-24.

[19] P. Qiu, Embracing the dropouts in single-cell RNA-seq analysis, Nat Commun 11(1) (2020) 1169.

[20] M. Andreatta, S.J. Carmona, UCell: Robust and scalable single-cell gene signature scoring, Comput Struct Biotechnol J 19 (2021) 3796-3798.

[21] X. Xie, Q. Shi, P. Wu, X. Zhang, H. Kambara, J. Su, H. Yu, S.Y. Park, R. Guo, Q. Ren, S. Zhang, Y. Xu, L.E. Silberstein, T. Cheng, F. Ma, C. Li, H.R. Luo, Single-cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection, Nat Immunol 21(9) (2020) 1119-1133.

[22] L. Aumailley, S. Bourassa, C. Gotti, A. Droit, M. Lebel, Vitamin C Differentially Impacts the Serum Proteome Profile in Female and Male Mice, J Proteome Res 20(11) (2021) 5036-5053.

[23] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res 43(7) (2015) e47.

[24] A. Chrysanthopoulou, I. Mitroulis, E. Apostolidou, S. Arelaki, D. Mikroulis, T. Konstantinidis, E. Sivridis, M. Koffa, A. Giatromanolaki, D.T. Boumpas, K. Ritis, K. Kambas, Neutrophil extracellular traps promote differentiation and function of fibroblasts, J Pathol 233(3) (2014) 294-307.

[25] A. Curaj, D. Schumacher, M. Rusu, M. Staudt, X. Li, S. Simsekyilmaz, V. Jankowski, J. Jankowski, A.R. Dumitrascu, D.J. Hausenloy, A. Schuh, E.A. Liehn, Neutrophils Modulate Fibroblast Function and Promote Healing and Scar Formation after Murine Myocardial Infarction, Int J Mol Sci 21(10) (2020).

[26] D.Y. Tzeng, T.F. Deuel, J.S. Huang, R.M. Senior, L.A. Boxer, R.L. Baehner, Platelet-derived growth factor promotes polymorphonuclear leukocyte activation, Blood 64(5) (1984) 1123-8.

[27] L.M. Pierini, M.A. Lawson, R.J. Eddy, B. Hendey, F.R. Maxfield, Oriented endocytic recycling of alpha5beta1 in motile neutrophils, Blood 95(8) (2000) 2471-80.

[28] J. Patten, K. Wang, Fibronectin in development and wound healing, Adv Drug Deliv Rev 170 (2021) 353-368.

[29] B. Selvarajah, I. Azuelos, M. Plate, D. Guillotin, E.J. Forty, G. Contento, H.V. Woodcock, M. Redding, A. Taylor, G. Brunori, P.F. Durrenberger, R. Ronzoni, A.D. Blanchard, P.F. Mercer, D. Anastasiou, R.C. Chambers, mTORC1 amplifies the ATF4-dependent de novo serine-glycine pathway to supply glycine during TGF-beta(1)-induced collagen biosynthesis, Sci Signal 12(582) (2019).

[30] G.S. Ashcroft, Sex differences in wound healing, Advances in Molecular and Cell Biology, Elsevier2004, pp. 321-328.

[31] B. Rønø, L.H. Engelholm, L.R. Lund, A. Hald, Gender Affects Skin Wound Healing in Plasminogen Deficient Mice, PLOS ONE 8(3) (2013) e59942.

[32] T. Sokka, S. Toloza, M. Cutolo, H. Kautiainen, H. Makinen, F. Gogus, V. Skakic, H. Badsha, T. Peets, A. Baranauskaite, P. Geher, I. Ujfalussy, F.N. Skopouli, M. Mavrommati, R. Alten, C. Pohl, J. Sibilia, A. Stancati, F. Salaffi, W. Romanowski, D. Zarowny-Wierzbinska, D. Henrohn, B. Bresnihan, P. Minnock, L.S. Knudsen, J.W. Jacobs, J. Calvo-Alen, J. Lazovskis, R. Pinheiro Gda, D. Karateev, D. Andersone, S. Rexhepi, Y. Yazici, T. Pincus, Q.-R. Group, Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study, Arthritis Res Ther 11(1) (2009) R7.

[33] D. Tanaka, T. Kagari, H. Doi, T. Shimozato, Essential role of neutrophils in anti-type II collagen antibody and lipopolysaccharide-induced arthritis, Immunology 119(2) (2006) 195-202.

[34] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc Natl Acad Sci U S A 102(43) (2005) 15545-50.

[35] T.F. Fuller, A. Ghazalpour, J.E. Aten, T.A. Drake, A.J. Lusis, S. Horvath, Weighted gene coexpression network analysis strategies applied to mouse weight, Mamm Genome 18(6-7) (2007) 463-72.

[36] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinformatics 9 (2008) 559.

[37] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16(5) (2012) 284-7.

[38] G. Yu, Q.Y. He, ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization, Mol Biosyst 12(2) (2016) 477-9.

Legends to figures:

Figure 1. Female neutrophils exhibit higher transcriptional expression of extra-cellular matrix related genes

(**A**) Experimental design of primary neutrophil transcriptomic datasets reanalyzed in this study [11, 12]. Panel created with BioRender.com. (**B**) Bubble plot showing GSEA enrichment results for ECM-related genes based on sex expression from neutrophil bulk RNA-seq (all terms are female-biased). See full analysis results in **Supplementary Table S2B**. (**C**) Dotplot of 'Ucell' scores of ECM-related genesets aggregated by sample of origin in single cell neutrophil RNA-seq dataset from [12]. For each geneset, statistical significance in a non-parametric Wilcoxon rank sum test between the distribution of UCell scores between female and male neutrophils is reported. (**D**) GSEA plot of collagen genes from neutrophil bulk RNA-seq (FDR < 1e-8, indicative of female neutrophil expression bias). (**E**) Heatmap displaying the expression of collagen genes (gene set identical as in C) from the bulk neutrophil RNA-seq.

Figure 2. WGCNA reveals that the top female-biased expression module is enriched for collagen and ECM-related genes.

(A) Module-trait relationships from neutrophil bulk RNA-seq WGCNA. Note that the Salmon module (bold) is the most biased for female neutrophils. Also see **Figure S1** and **Supplementary Table S3.** (B) Heatmap of sample-wise module eigengenes from bulk neutrophil RNA-seq WGCNA analysis. (C) Dotplot of significantly enriched gene sets associated to the Salmon module, using hypergeometric enrichment analysis with Gene Ontology [GO] terms (FDR < 0.05). Also see full results in **Supplementary Table S3**. (D) Gene Concept network for GO enrichment analysis related to the Salmon cluster, derived from the same analysis as (C).

Legends to Supplementary figures:

Supplementary Figure S1. Analysis of ECM-related gene expression in bulk neutrophil RNA-seq, single-cell neutrophil RNA-seq and serum proteomics.

(A) Bubble plot showing Fisher's exact test enrichment results for overlap of significantly female-biased genes in neutrophil bulk transcriptomes (DESeq2 FDR < 5%) and ECM-related genes. See full results in **Supplementary Table S2C**. (B) Dotplot of 'Ucell' scores of ECM-related genesets aggregated by neutrophil maturation subset (as defined by [21]) in single-cell neutrophil RNA-seq dataset from [12]. (C) Multidimensional scaling analysis for mouse serum proteomics. MDS: Multidimensional Scaling. (D) Heatmap of significant (Limma FDR < 5%) sex-dimorphic proteins in mouse serum. Also see **Supplementary Table S2D**. (E) Bubble plot showing Fisher's exact test enrichment results for overlap of significantly female-biased proteins in serum proteomics (Limma FDR < 5%) and ECM-related proteins. See full results in **Supplementary Table S2E**.

Supplementary Figure S2. WGCNA module cluster dendrogram and salmon module expression profile.

(A) WGCNA module cluster dendrogram from bulk neutrophil RNA-seq. (B) Heatmap of WGCNA Salmon module gene expression from bulk neutrophil RNA-seq.

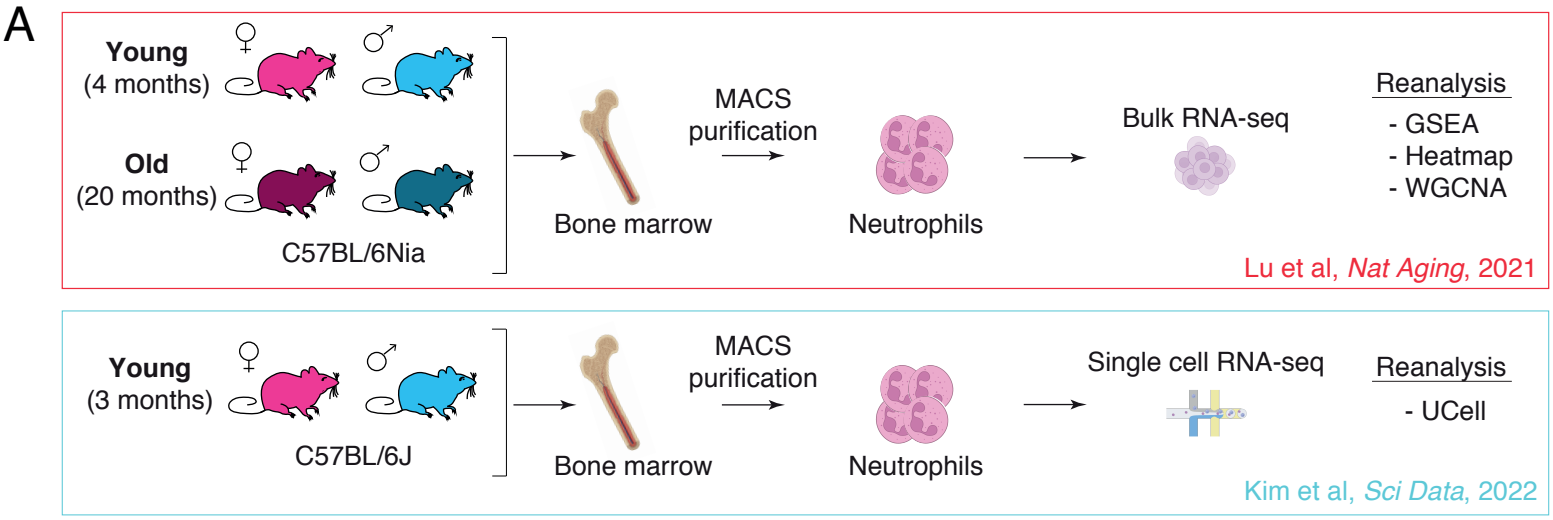
Inventory of Supplementary tables:

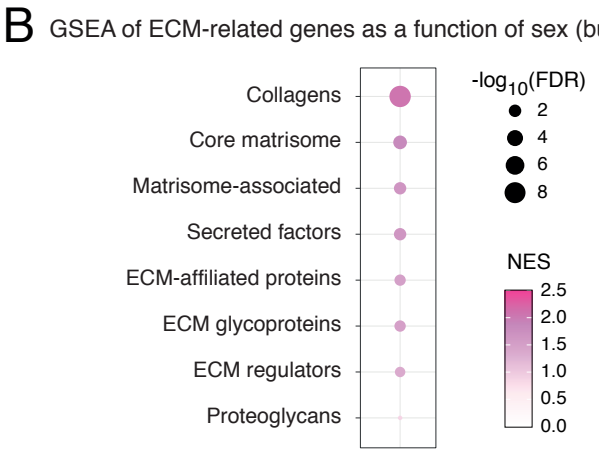
Supplementary Table S1: Curated Gene Lists for ECM-related genes.

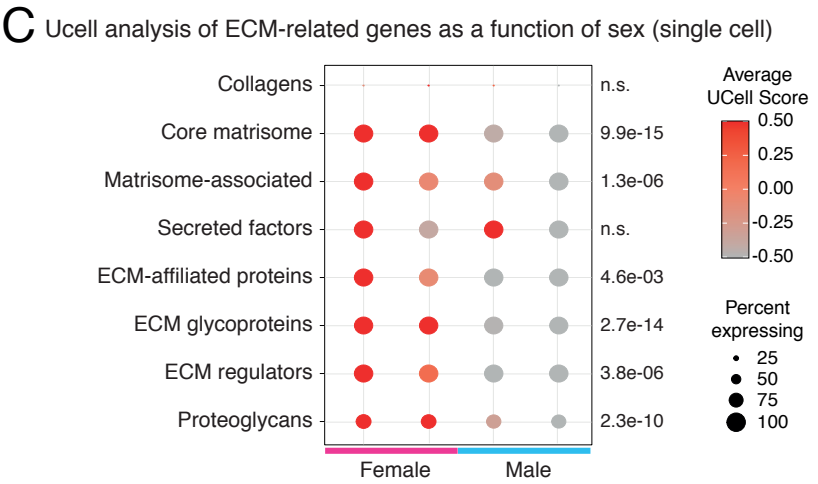
Supplementary Table S2: Enrichment analysis results for ECM-related gene lists. (A) GSEA results for ECM-related gene lists on bulk neutrophil RNA-seq, as a function of age (sex as covariate). Note that no ECM gene set passed significance thresholds for age-related expression. (B) GSEA results for ECM-related gene lists on bulk neutrophil RNA-seq in female vs. male neutrophils (age as covariate). (C) Analysis of significance of overlap for female-biased neutrophil genes by RNA-seq (DEseq2 FDR < 0.05) for ECM related gene list. (D) Limma differential expression analysis of sex-biased proteins in C57BL/6 serum proteomics from [22] (FDR < 0.05). (E) Analysis of significance of overlap for female-biased serum proteins by proteomics (Limma FDR < 0.05) for ECM-related gene list.

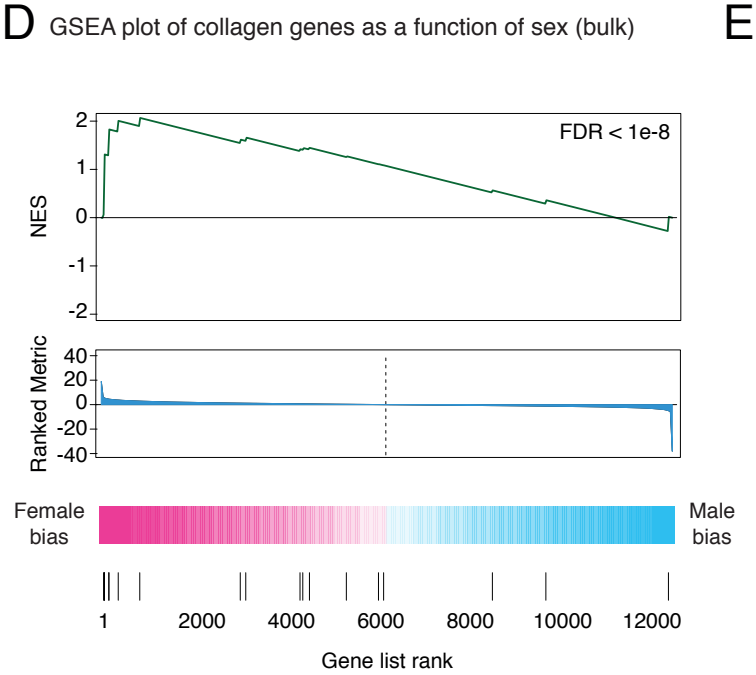
Supplementary Table S3: WGCNA related analysis with module membership, and Gene Ontology and Reactome hypergeometric enrichment analysis of Salmon and magenta modules.

Figure 1


A
Young
(4 months)
Old
(20 months)
C57BL/6Nia
Bone marrow
MACS
purification
Neutrophils
Bulk RNA-seq
Reanalysis
- GSEA
- Heatmap
- WGCNA
Lu et al, Nat Aging, 2021
Young
(3 months)
C57BL/6J
Bone marrow
MACS
purification
Neutrophils
Single cell RNA-seq
Reanalysis
- UCell
Kim et al, Sci Data, 2022


B GSEA of ECM-related genes as a function of sex (bulk)
Collagens
Core matrisome
Matrisome-associated
Secreted factors
ECM-affiliated proteins
ECM glycoproteins
ECM regulators
Proteoglycans
-log10(FDR)
2
4
6
8
NES
2.5
2.0
1.5
1.0
0.5
0.0


C Ucell analysis of ECM-related genes as a function of sex (single cell)
Collagens
Core matrisome
Matrisome-associated
Secreted factors
ECM-affiliated proteins
ECM glycoproteins
ECM regulators
Proteoglycans
n.s.
9.9e-15
1.3e-06
n.s.
4.6e-03
2.7e-14
3.8e-06
2.3e-10
Female
Male
Average
UCell Score
0.50
0.25
0.00
-0.25
-0.50
Percent
expressing
25
50
75
100


D GSEA plot of collagen genes as a function of sex (bulk)
FDR < 1e-8
NES
Ranked Metric
Female
bias
Male
bias
1
2000
4000
6000
8000
10000
12000
Gene list rank

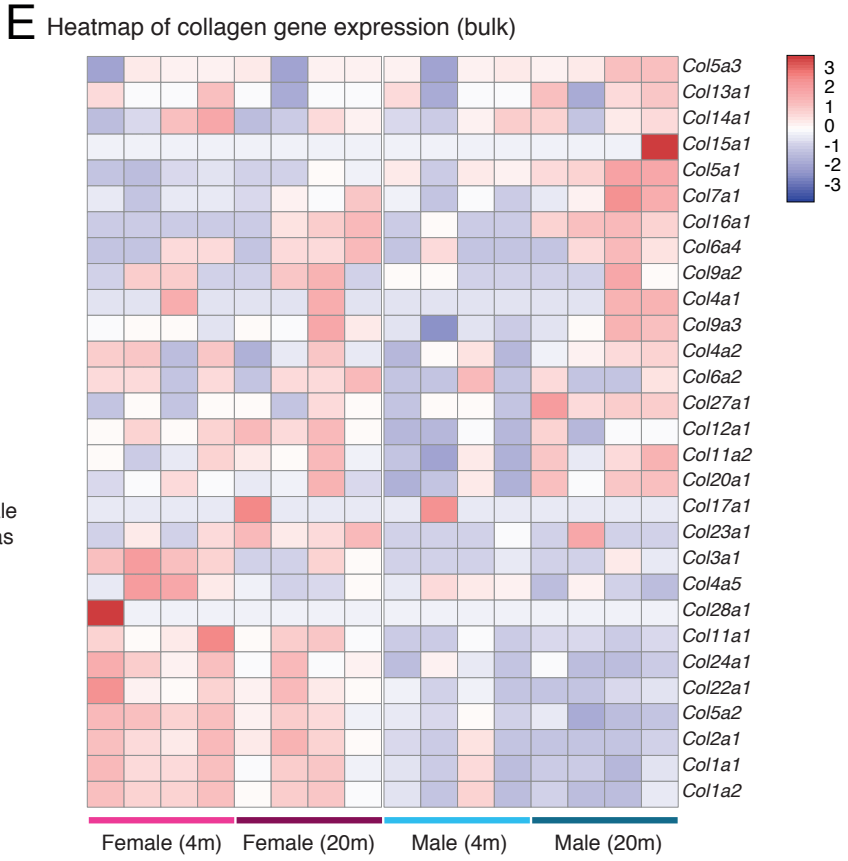
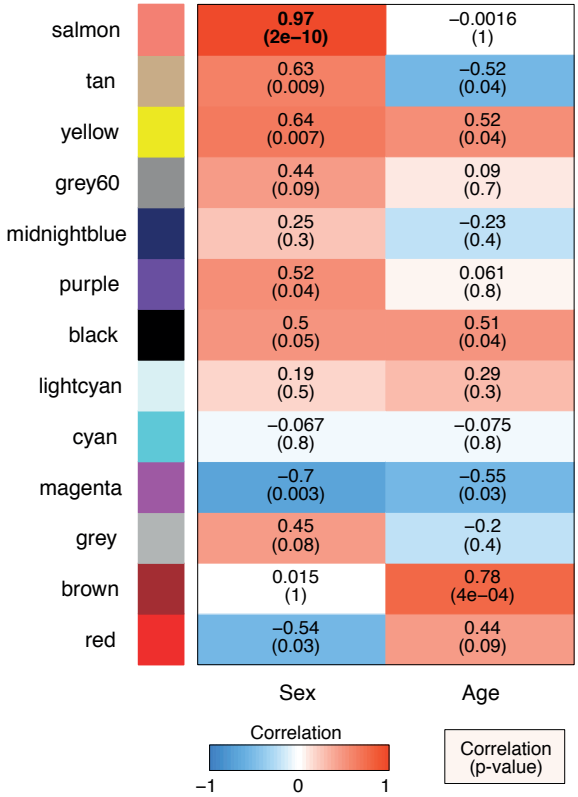
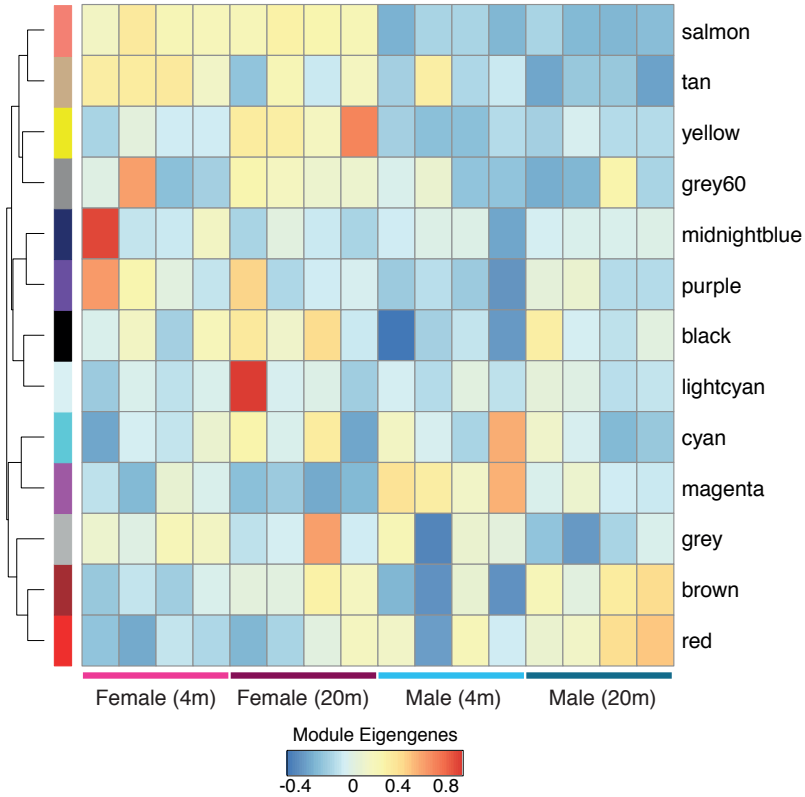

E Heatmap of collagen gene expression (bulk)
Col5a3
Col13a1
Col14a1
Col15a1
Col5a1
Col7a1
Col16a1
Col6a4
Col9a2
Col4a1
Col9a3
Col4a2
Col6a2
Col27a1
Col12a1
Col11a2
Col20a1
Col17a1
Col23a1
Col3a1
Col4a5
Col28a1
Col11a1
Col24a1
Col22a1
Col5a2
Col2a1
Col1a1
Col1a2
Female (4m)
Female (20m)
Male (4m)
Male (20m)

Figure 2

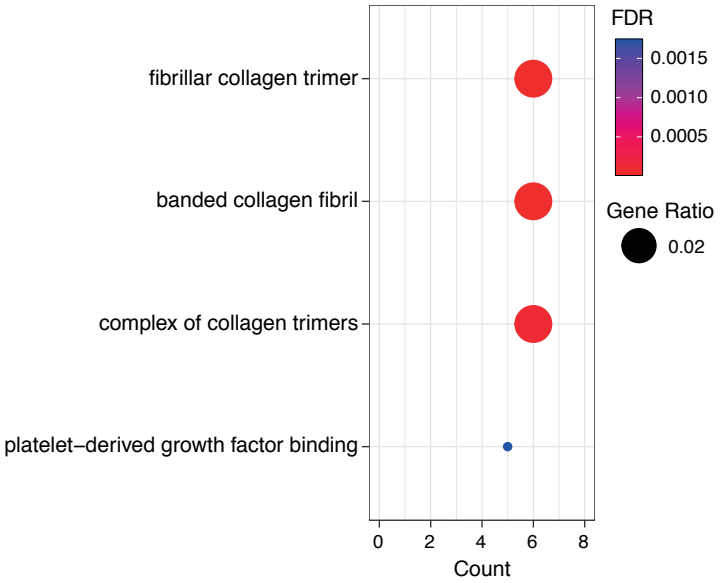
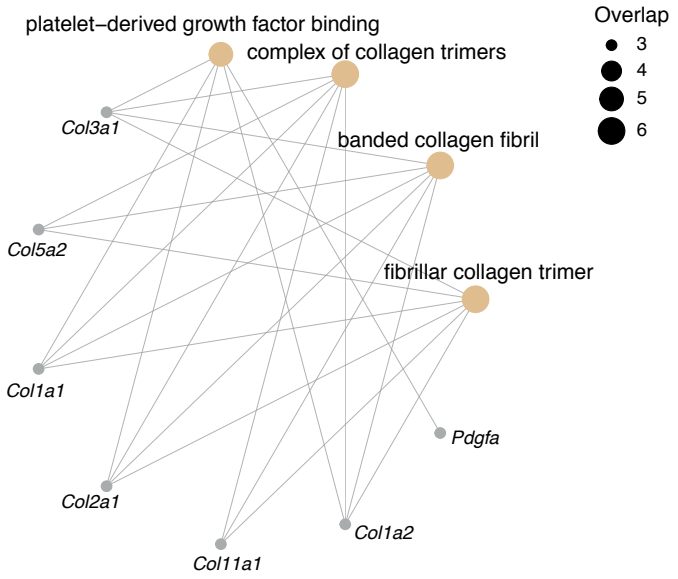
A Module-trait relationships from WGCNA (bulk)



B Heatmap of module eigengenes by sample from WGCNA (bulk)



C Salmon cluster hypergeometric enrichments (Gene Ontology)



D Salmon cluster gene-concept network (Gene Ontology)