Supplementary Material

Methods

Tissue sample preparation and spatial transcriptomics sequencing

Fresh frozen tissue samples were embedded in OCT compound and cryosectioned. Tissue sections were mounted onto Visium capture slides, fixed, and stained with hematoxylin and eosin (H&E) for histological imaging. Following imaging, tissue permeabilization was performed to release mRNA molecules, which were captured by spatially barcoded oligonucleotides present on the slide surface. For further steps, the standard 10x Genomics Visium Spatial Gene Expression workflow pipeline was followed.

Downstream computational analysis of spatial transcriptomics data

All the spatial transcriptomics analyses were conducted in R v4.4.0 using a pipeline combination of Seurat v5.4 (65) and SPATA2 v3.1.3 (66) packages. The primary input for both pipelines consisted of the output directories generated by Cell Ranger (10x Genomics), including the filtered feature-barcode matrices, spatial image data and associated metadata.

The Seurat analysis pipeline was employed for unsupervised clustering, dimensionality reduction, normalization (via SCTransform), differential gene expression analysis and deconvolution of the spatial transcriptomics data using Steele et al.'s Skin Atlas scRNA reference (30). These steps enabled us to identify transcriptionally distinct spatial domains in an unbiased manner, perform comparisons between samples and assess the transcriptional similarity of the fibroblast subpopulations from the Skin Atlas with the spots of our spatial slides.

To facilitate the cluster recognition across all samples, we performed integration of the 14 individual slides using Seurat's SCTransform v2-based integration workflow (67) by selecting a list of features for the integration and finding integration anchors.

After computing the k = 20 nearest neighbors (KNN) of the dataset, we proceeded with the identification of functional cell clusters. Our chosen parameters for the clustering were a resolution of 0.5 and the Louvain algorithm with multilevel refinement. The deconvolution is based on Seurat v3's anchor-based integration workflow.

Differential expression analysis was performed using FindAllMarkers() and FindMarkers() functions implemented in the Seurat. Spatial transcriptomic spots were grouped according to experimental condition, and pairwise comparisons were conducted using the Wilcoxon Ranked Sum test.

SPATA2 enabled us to perform supervised clustering for differential gene expression analysis and Gene Set Enrichment Analysis. We also used SPATA2 to define a fibrosis development trajectory, followed by Spatial Trajectory Screening (STS) to discover genes that show a non-random expression pattern along the fibrotic axis.

We used clusterProfiler v4.12.6 and the MSigDB, Reactome and Molecular Function (from gene Ontologies) collections for Gene Set Enrichment Analysis of intersample differential expression analysis results (68–72).

Robust Cell Type Decomposition (RCTD), implemented in the spacexr R package (75), was used to infer the relative contribution of different cell types to the transcriptomic signal of each spatial spot. As a reference, we used the annotated single-cell RNA-sequencing dataset of localized scleroderma skin published by Hutchins et al. (2024) under the accession number GSE264508 (76). RCTD allowed the cell-type composition of individual spots to be estimated and the spatial distribution of major cell populations across the tissue to be characterized.

Identification of compounds predicted to oppose morphea-associated transcriptional programs

This analysis was adapted from a previously described pathway-correlation framework in which disease-associated pathway enrichment profiles were compared with perturbational

signatures from the CMap resource (73). To prioritize compounds with the potential to reverse morphea-associated transcriptional alterations, we performed a pathway-level perturbational analysis using the Broad Institute Connectivity Map (CMap) L1000 resource. We used replicate-collapsed level 5 perturbational signatures from the L1000 dataset and restricted the analysis to small-molecule perturbagens. For each spatial transcriptomic comparison, genes were ranked by log fold change, and KEGG pathway enrichment analysis was performed using fgsea, yielding normalized enrichment scores (NES) for each pathway. Pathway enrichment was calculated analogously for the L1000 perturbational signatures. We then quantified the similarity between morphea-associated and perturbation-associated pathway profiles by calculating Pearson correlation coefficients between the corresponding KEGG NES vectors as described in (74). Negative correlations were interpreted as indicating compounds with pathway-level effects opposing the morphea signature, whereas positive correlations indicated concordant effects. Since we focused on drug repurposing, the results were filtered for approved, clinical, preclinical and non-clinical compounds, and not on compounds only available in screening libraries. Compound annotations were reviewed using PubChem and manual literature searches as a downstream curation step. PubChem was used only for compound annotation and not as part of the statistical prioritization procedure.

Given the redundancy of the Connectivity Map resource across experimental conditions, perturbational instances were aggregated at the compound level. For each compound within each disease signature, summary statistics were computed across all available contexts, including: number of supporting signatures, number of distinct cell lines, number of treatment, durations and doses, median connectivity score, minimum connectivity score, proportion of negative connectivity scores.

Median connectivity score was used as the primary measure of effect size, providing an estimate of the compound’s overall tendency to reverse the disease-associated transcriptional signature.

To identify compounds predicted to reverse disease-associated transcriptional signature, only compounds with negative median connectivity scores were considered. A minimum number of supporting perturbational instances (≥2 signatures per compound) was required for robustness. Compounds were then ranked within each disease signature based on a combination of median connectivity score, minimum score, and level of experimental support.

Top candidate compounds were defined as the highest-ranking compounds (usually top 15 per disease, unless insufficient compounds with negative connectivity score per comparison) satisfying these criteria. This selection was performed independently for each disease signature. Aiming to identify compounds with consistent reversal effects across multiple comparisons, overlap analysis was performed on the top-ranked compounds from each disease.

Second harmonic generation

Second harmonic generation (SHG) imaging was performed on 8-µm cryosections of a spindle-shaped skin biopsy using an upright multiphoton microscope (TCS SP8 MP-OPO, Leica Microsystems, Wetzlar, Germany) equipped with a Ti laser (Chameleon Vision II, Coherent). Fibrillar collagen was visualized by label-free SHG imaging to assess differences in collagen distribution and spatial organization across the tissue. The spindle-shaped biopsy provided a continuous tissue section encompassing the clinically uninvolved, inflammatory, and fibrotic areas, enabling direct visualization and comparison of collagen architecture along the spatial transition between these disease zones. Images were acquired using consistent acquisition parameters and LAS X software (Leica Microsystems). SHG imaging was performed at the CECAD Imaging Facility, University of Cologne.

Immunofluorescence

Skin cryosections (8 µm) were fixed in methanol/acetone (1:1) for 5 min at −20 °C and blocked for 1 h at room temperature in PBS containing 1% FCS. Sections were then incubated

overnight at 4 °C with primary antibodies diluted in blocking solution: rabbit polyclonal anti-COMP (1:500) (82), rabbit polyclonal anti-tenascin X (1:500; kind gift of Dr. Manuel Koch, University of Cologne, Germany) and goat polyclonal anti-collagen I (1:500; AB758, Chemicon/Merck Millipore, Burlington, MA, USA). After three 10-min washes in PBS, sections were incubated for 1 h at room temperature with the appropriate highly cross-adsorbed secondary antibodies conjugated to AlexaFluor 488 or 555 (1:400; Thermo Fisher Scientific, Waltham, MA, USA). Nuclei were counterstained with DAPI (0.1 µg/ml; Sigma-Aldrich, St. Louis, MO, USA). Sections were washed as above, mounted in Dako Fluorescence Mounting Medium (Dako, Glostrup, Denmark) and imaged on a Leica DMi8 microscope. Images were processed with ImageJ (NIH, Bethesda, MD, USA).

Supplementary Tables' Legends

Supplementary Table 1: Clinical characteristics of the study cohort and biopsy specimens.

Demographic and clinical characteristics, biopsy sites, sampling strategies, disease duration, and treatment history of patients with plaque morphea (PM; n = 2), generalized morphea (GM; n = 4), and eosinophilic fasciitis (EF; n = 3), as well as healthy controls (n = 2). For PM_1, a single spindle-shaped biopsy spanning clinically unaffected, inflammatory, and fibrotic regions was analyzed in two independent sections (PM_1a and PM_1b). For PM_2, three separate punch biopsies were obtained from clinically unaffected (PM_2a), inflammatory (PM_2b), and fibrotic (PM_2c) regions. All other individuals contributed a single biopsy. Previous treatments and the most recent treatment prior to biopsy are indicated.

Supplementary Table 2: Summary of quality control metrics for all Visium spatial transcriptomics samples. Reported metrics include sequencing depth, spot coverage, barcode and UMI validity, sequencing quality scores, mapping statistics, and transcript detection metrics generated during Space Ranger processing. For more information about the full metrics please refer to: https://www.10xgenomics.com/support/software/space-ranger/latest/analysis/outputs/space-ranger-web-summary

Supplementary Table 3: Expression dynamics of all the 39 trajectory-associated genes with an FDR < 0.05 identified with SPATA2. The table includes a summary of the model fitting metrics of every gene's expression gradient. Quality of the fit is reported using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The lower the values, the better the quality of each fit.

Supplementary table 1: Characterization of patients and of biopsies

Subtype	Patient ID	Number of biopsies	Site of biopsy	Gender	Age [years]	Disease duration	Clinical Findings	Previous therapy	Latest treatment prior to biopsy (discontinued ≥ 3 weeks)
Plaque Morphea	**PM_1***	One spindle biopsy spanning from unaffected to inflammatory to fibrotic areas of the lesion (two independent analyses **PM_1a** and **PM_1b**)	**para-vertebral**	f	69	~16 years	multiple confluent plaques with sclerotic center and lilac ring	UVA1, topical	None
	PM_2*	Three punch biopsies from the unaffected (**PM_2a**), inflammatory (**PM_2b**) and fibrotic (**PM_2c**) areas of the lesion	**hip / lower abdomen**	m	53	~2 years	singular plaque with sclerotic center and lilac ring	topical	None
Generalized Morphea	**GM_1**	One unique punch biopsy	**abdomen**	f	75	~6 months	large confluent areas of thick, hardened skin on the trunk, & extremities	none	None
	GM_2	One unique punch biopsy	**lower abdomen**	m	62	~4 months		none	None
	GM_3	One unique punch biopsy	**abdomen**	m	80	~10 years		MTX, MFA, UVA1	None
	GM_4	One unique punch biopsy	**abdomen**	m	61	~2 years		none	None
Eosinophilic Fasciitis	**EF_1**	One unique punch biopsy	**lower leg**	f	42	~19 years	thick, hardened skin of both lower legs	MTX	None
	EF_2	One unique punch biopsy	**lower leg**	m	54	~9 months	thick, hardened skin and swelling of both lower legs	none	None
	EF_3	One unique punch biopsy	**lower leg**	f	51	~7 yrs	thick, hardened skin and swelling of both lower legs	MTX, SULF for 3 years, HCQ	HCQ
Control	**Ctrl_1**	One unique punch biopsy	**lower leg**	f	56	--	--	--	--
	Ctrl_2	One unique punch biopsy	**abdomen**	m	53	--	--	--	--

Abbreviations: MTX = methotrexate; SULF = sulfasalazine; HCQ = hydroxychloroquine; ANA = anti-nuclear antibodies; CRP = C-reactive protein; MFA = mycophenolic acid; * In the case of patients from which multiple biopsies were collected, a letter was added to each biopsy's ID to be able to distinguish them from each other.

Supplementary Table 2: Quality control metrics of all samples.

Name	Number of Spots Under Tissue	Number of Reads	Mean Reads per Spot	Mean Reads Under Tissue per Spot	Fraction of Spots Under Tissue	Median Genes per Spot	Median UMI Counts per Spot	Valid Barcodes	Valid UMIs	Sequencing Saturation	Q30 Bases in Barcode	Q30 Bases in RNA Read	Q30 Bases in UMI	Reads Mapped to Genome	Reads Mapped Confidently to Genome	Reads Mapped Confidently to Intergenic Regions	Reads Mapped Confidently to Intronic Regions	Reads Mapped Confidently to Exonic Regions	Reads Mapped Confidently to Transcriptome	Reads Mapped Antisense to Gene	Fraction Reads in Spots Under Tissue	Total Genes Detected
Ctrl_1	968	89101935	92047.45351	81390.59711	0.193910256	534	700	0.986259872	0.999364223	0.970844533	0.937790822	0.922650853	0.927833773	0.564340887	0.529967694	0.015429194	0.013687458	0.500851042	0.4898755	0.002807481	0.881755409	17184
Ctrl_2	836	86757925	103777.4222	96861.98325	0.167467949	1318.5	2516	0.988223266	0.998604957	0.953840311	0.956388149	0.937740203	0.950918364	0.905207588	0.873270943	0.01322874	0.018544715	0.841497489	0.823189789	0.003764636	0.946582227	19461
EF_1	2084	207684868	99656.84645	82677.92035	0.417467949	925	1366	0.950777646	0.999633637	0.968807149	0.944607675	0.934369246	0.943008761	0.90625122	0.784556899	0.025542405	0.023121083	0.735893411	0.719626328	0.004112986	0.914324653	18999
EF_2	1297	175405806	135239.6345	120697.1164	0.259815705	914	1553	0.987734175	0.998636083	0.973616261	0.956567678	0.939091466	0.952223673	0.892554594	0.867239497	0.015801204	0.015953839	0.835484454	0.818774095	0.003900373	0.906840108	19483
EF_3	2077	222560197	107154.6447	95907.06066	0.416065705	634	978	0.98819133	0.99859966	0.975350096	0.956988382	0.933541072	0.951863312	0.78485263	0.739245302	0.011023476	0.018294835	0.709926991	0.694804085	0.0042145	0.896323121	18754
GM_1	1862	228294341	122607.0575	113029.6729	0.372996795	359	537	0.987663531	0.998677374	0.988388959	0.957865319	0.940680842	0.953494094	0.897841896	0.838751312	0.011734399	0.022787398	0.804229514	0.789815359	0.003292732	0.930698306	17885
GM_2	741	74970589	101174.8839	93137.26316	0.1484375	197	242	0.987686478	0.999622772	0.984718212	0.940715214	0.93713964	0.934367117	0.896028121	0.86304516	0.014994813	0.015172657	0.83287769	0.817239171	0.003568866	0.934107273	16099
GM_3	1281	162442124	126808.84	115647.9571	0.256610577	880	1513	0.986228566	0.999664133	0.976933577	0.943271101	0.934901657	0.936618478	0.911504931	0.890299951	0.014094897	0.015467804	0.860737249	0.843934305	0.003454276	0.924889498	19080
GM_4	569	48489104	85218.10896	76936.34974	0.113982372	547	821	0.987233854	0.999637217	0.97521546	0.940948672	0.937340305	0.933929938	0.900768428	0.87712425	0.015993284	0.015783587	0.845347379	0.829810219	0.003493238	0.915278288	16771
PM_1a	1292	119073353	92162.03793	72685.72523	0.258814103	1652	3125.5	0.958587057	0.996954432	0.908848587	0.966971307	0.953523209	0.966302911	0.920808117	0.885073304	0.013125472	0.015706033	0.856241799	0.835375754	0.005643588	0.825011182	20525
PM_1b	1471	129202000	87832.76683	74296.04827	0.294671474	1374	2388	0.962957346	0.998087607	0.925035629	0.967485596	0.95477812	0.96672976	0.942135354	0.903710616	0.012644518	0.015988081	0.875078017	0.854855149	0.00578176	0.880153915	20687
PM_2a	717	101380867	141395.9093	119312.2845	0.143629808	275	360	0.980476533	0.998872312	0.983676663	0.963349336	0.948096047	0.957345933	0.872931329	0.840168727	0.021647882	0.02014678	0.798374066	0.781053184	0.005221538	0.866185188	16925
PM_2b	893	123558448	138363.3236	114297.3471	0.178886218	358	464	0.979629632	0.99745654	0.976080678	0.960681342	0.930739959	0.953147285	0.631779812	0.582149737	0.019522315	0.030338452	0.53228897	0.519554357	0.00454907	0.832670247	17089
PM_2c	586	97581797	166521.8379	132504.8549	0.117387821	373	536	0.980162407	0.998109699	0.979022442	0.962565542	0.947619531	0.958210695	0.829420276	0.786316335	0.023097433	0.019757445	0.743461457	0.728584758	0.005145581	0.815602833	16267

Supplementary Table 3: Genes and information relating to expression pattern trajectory.

variables	models	mae	rmse
AGAP3	ascending_late	0.19977	0.223481
ARPC1B	ascending_linear	0.257669	0.344779
C3	descending_linear	0.200681	0.255663
CD248	descending_linear	0.12402	0.163715
CD81	descending_linear	0.271511	0.339919
CDC34	ascending_late	0.165053	0.196198
CHCHD5	ascending_late	0.177248	0.211975
CNN3	ascending_late	0.178346	0.213298
COL1A1	descending_linear	0.185134	0.214484
COL3A1	descending_linear	0.19297	0.251899
CSNK1A1	ascending_late	0.182442	0.220081
CSRP1	descending_linear	0.288437	0.37452
FBLN1	descending_linear	0.159944	0.206326
GPAA1	ascending_late	0.158776	0.181935
HBB	ascending_late	0.129397	0.153641

HRAS	ascending_gradual	0.21274	0.249168
IGLC1	peak_gradual	0.23088	0.273032
KRT14	ascending_linear	0.139233	0.182772
LAD1	ascending_gradual	0.194255	0.239369
LGALS7	ascending_gradual	0.099368	0.121779
LGALS7B	ascending_gradual	0.094847	0.113391
MRPS34	ascending_late	0.194745	0.220051
NAA38	ascending_gradual	0.250253	0.280116
PERP	ascending_gradual	0.232109	0.267839
PSMA1	ascending_late	0.174415	0.209889
PSMB1	ascending_late	0.206252	0.238738
PSMB5	ascending_late	0.138804	0.175302
RETREG3	ascending_late	0.15006	0.185589
RPL11	peak_gradual	0.309343	0.39842
RPL18	ascending_linear	0.135918	0.162922
RPL32	peak_gradual	0.278208	0.349348
RPS3	ascending_linear	0.279726	0.328052

SBF1	ascending_late	0.197786	0.230291
TMUB2	ascending_late	0.144025	0.177447
TNXB	descending_linear	0.144662	0.1713
TRAF7	ascending_late	0.145934	0.173781
TSTA3	ascending_late	0.140809	0.157944
VOPP1	ascending_late	0.186995	0.21976
ZDHHC12	ascending_late	0.164526	0.185884

Supplementary Figures' Legends

Supplementary Figure 1: Quality control plots of each sample containing their respective H&E-stained image of the fresh-frozen tissue mounted on the spatial transcriptomics slide, their spatial distribution of total transcript counts (n Count), of detected genes (n Feature), and the percentage of transcripts mapping to mitochondrial (percentage.mt) per spot. Warmer colors indicate higher amount of transcript counts, higher number of detected genes, and higher mitochondrial gene content, respectively. Violin plots show the distributions of total transcript counts (left), detected genes (middle), and mitochondrial transcript percentage (right) across spots. Corresponding sample names are at the upper left corner of each panel.

Supplementary Figure 2: Batch effect-corrected unsupervised clustering data of each sample. Each sample's name is on top the corresponding sample's plot. On the right side of each plot a legend indicates the color mapping between each cluster's number and color.

Supplementary Figure 3: Characterization of the ECM in PM and HC; (A) Second harmonic generation (SHG) imaging was used to assess fibrillar collagen organization and distribution across PM_1a. (B) H&E staining of FFPE tissue sections was performed to illustrate the histomorphological differences between the sclerotic core, inflammatory border, clinically uninvolved regions, and healthy control skin. (C) Immunofluorescence staining for collagen I, COMP and tenascin-X was performed to visualize the spatial distribution of extracellular matrix components across these regions in PM_1a.

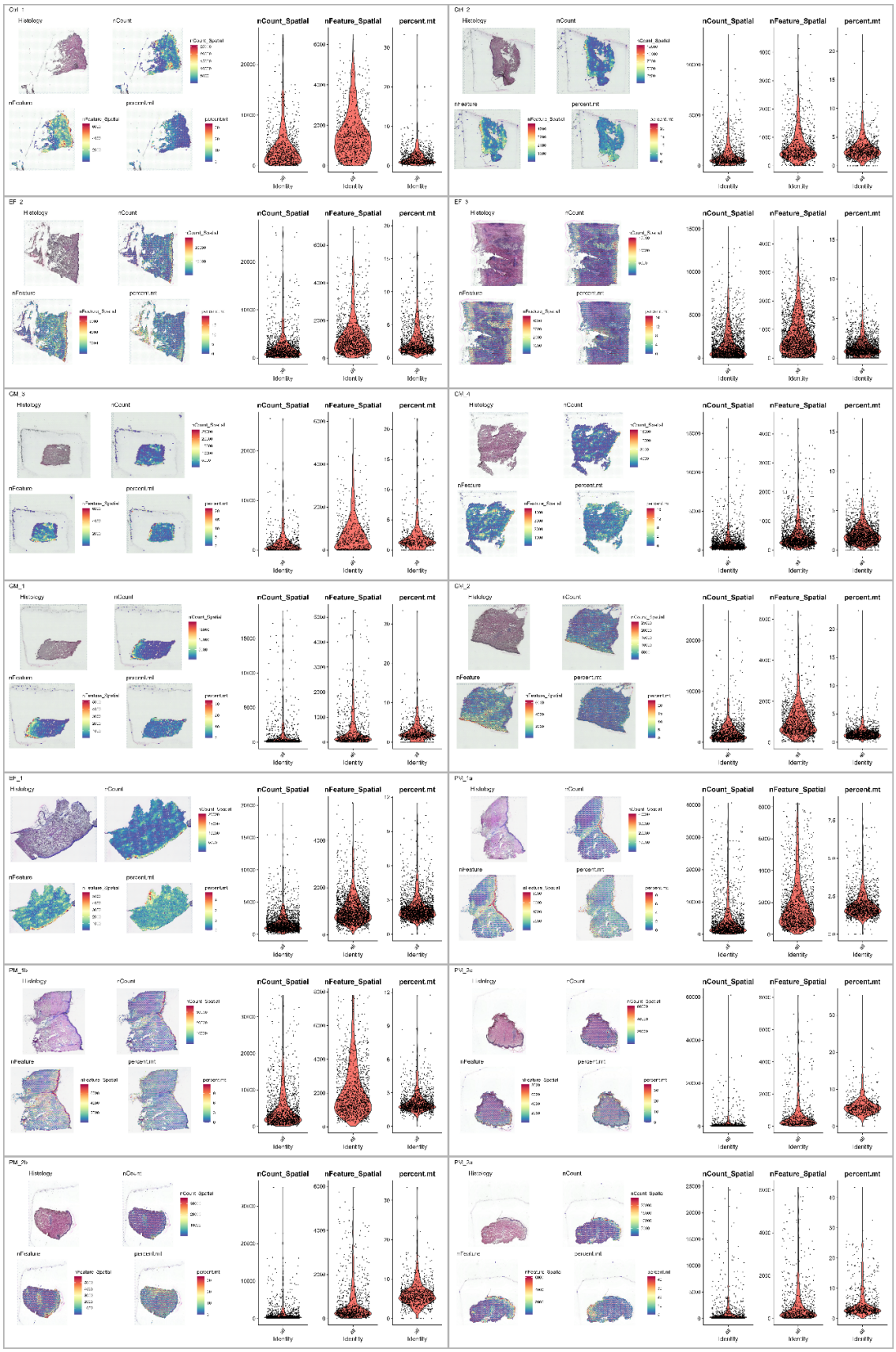
Supplementary Figure 4: On the left are Enrichr gene set enrichment analysis (GSEA) results for the MSigDB Hallmark collection, Reactome pathways, and the Gene Ontology Molecular Function category, sorted by descending combined score (defined as the p-value multiplied by the odds ratio). On the right, genes exhibiting a non-random expression gradient along the fibrotic trajectory are shown: genes with FDR < 0.1 are enclosed in a black rectangle, while those with FDR < 0.05 are enclosed in a blue rectangle. For the GSEA, only the 77 genes with FDR < 0.1 were tested against a background consisting of all genes detected in the PM_1a sample.

Supplementary Figure 5: Robust cell type decomposition of every sample, analogous to the one shown in Figure 6B. All cell types included in the scRNA reference used are present in every panel of its corresponding sample. Sample names are above each panel.

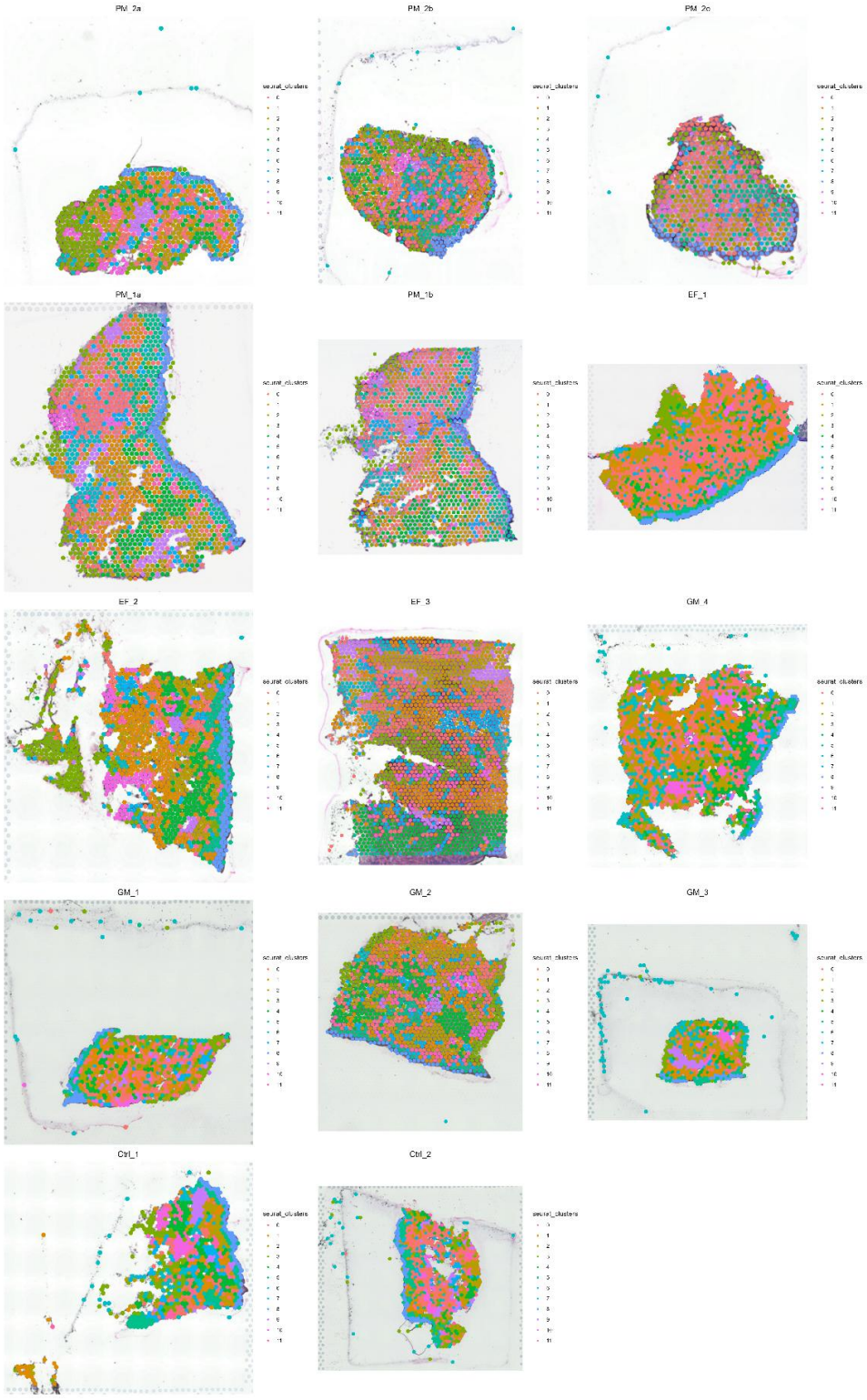
Supplementary Figure 6: Spatial deconvolution plots for each sample and each fibroblast subpopulation, analogous to those shown in Figure 6C. All cell types included in the scRNA reference used are present in every panel of its corresponding sample. Corresponding sample names are at the top of every panel.

Supplementary Figure 7: Predicted localization of superficial fibroblasts, FRC-like fibroblasts and inflammatory myofibroblasts subpopulations. Analysis of the spatial transcriptomics data for specific fibroblast subsets is based on the recently published single cell RNA atlas (30). Each panel represents the predicted spatial distribution of selected fibroblast subsets (F1:superficial, F3:FRC-like; F6: inflammatory myofibroblast) from the reference atlas concentrating on the samples with eosinophilic fasciitis (EF_1; EF_2, EF_3) and generalized morphea (GM_1, GM_2, GM_3 and GM_4). Color intensity reflects the prediction score for each spot. Higher scores indicate stronger transcriptomic similarity to the reference cell type. Prediction scores are relative confidence measures, and do not sum to 100% across cell types; a high score indicates that the spot's expression pattern is strongly supported for that cell type, while low scores indicate weak or no support. The color scale is panel-specific, spanning the minimum to maximum prediction score observed for that cell type.

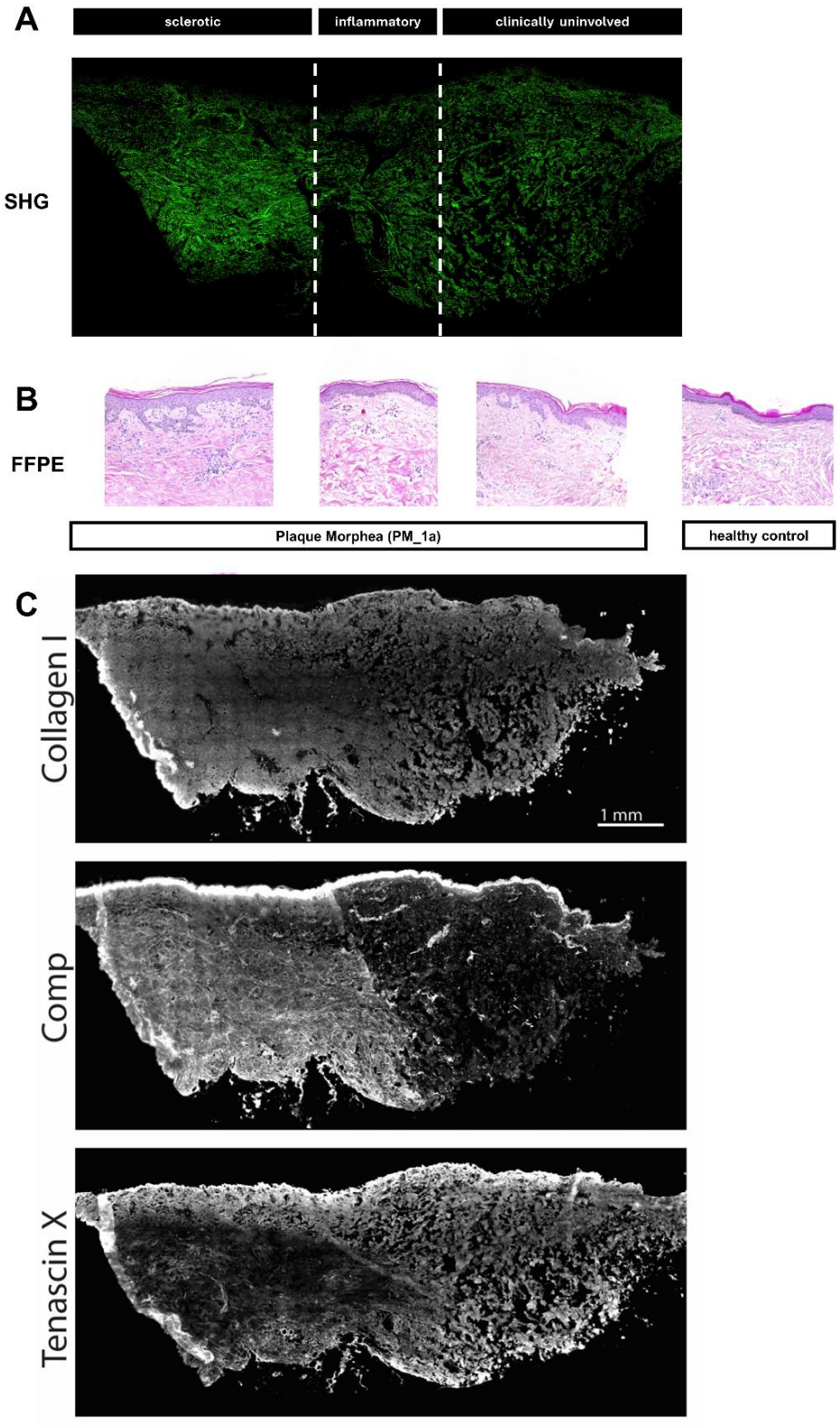
Supplementary Figure 1



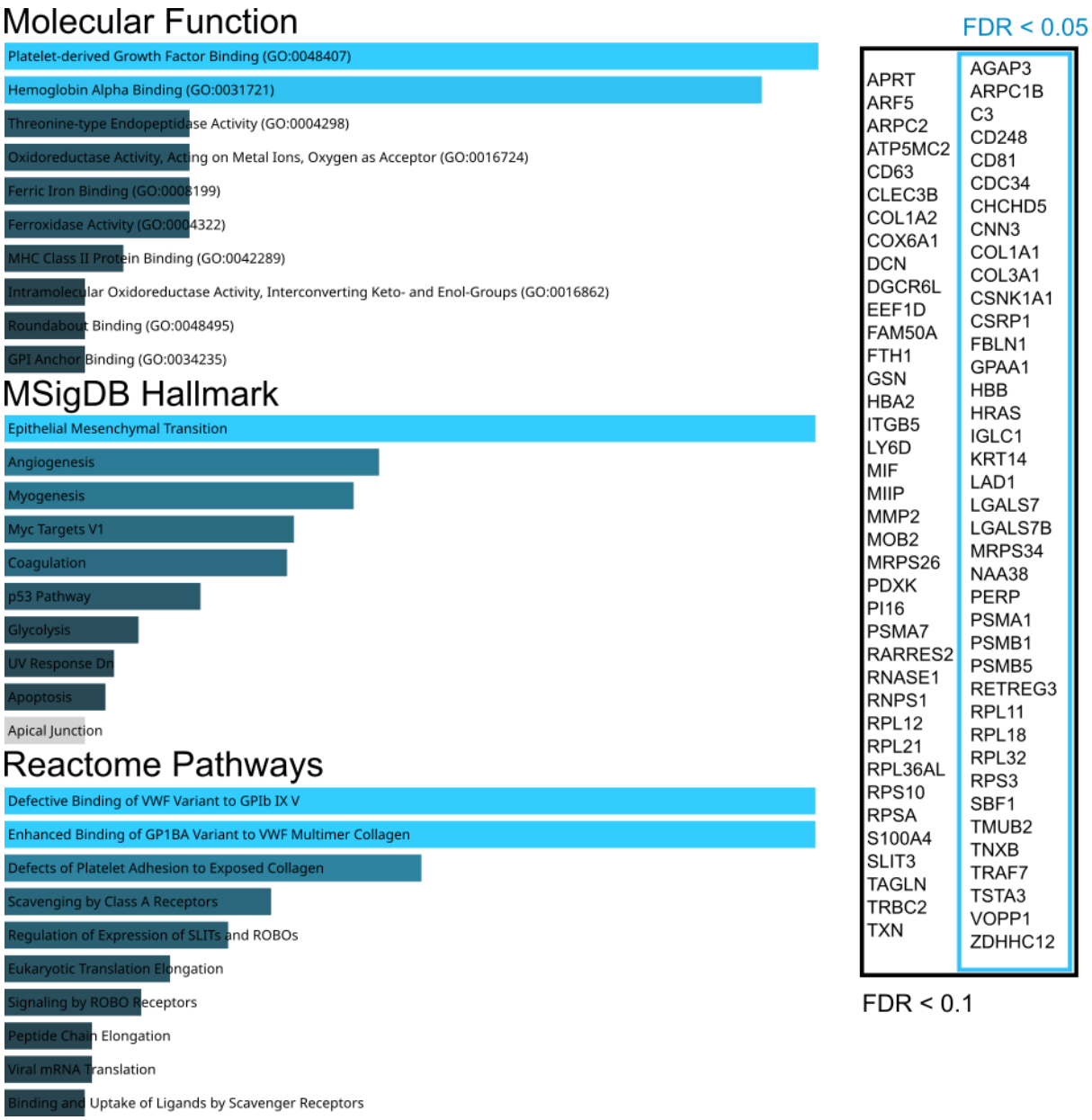
Supplementary Figure 2



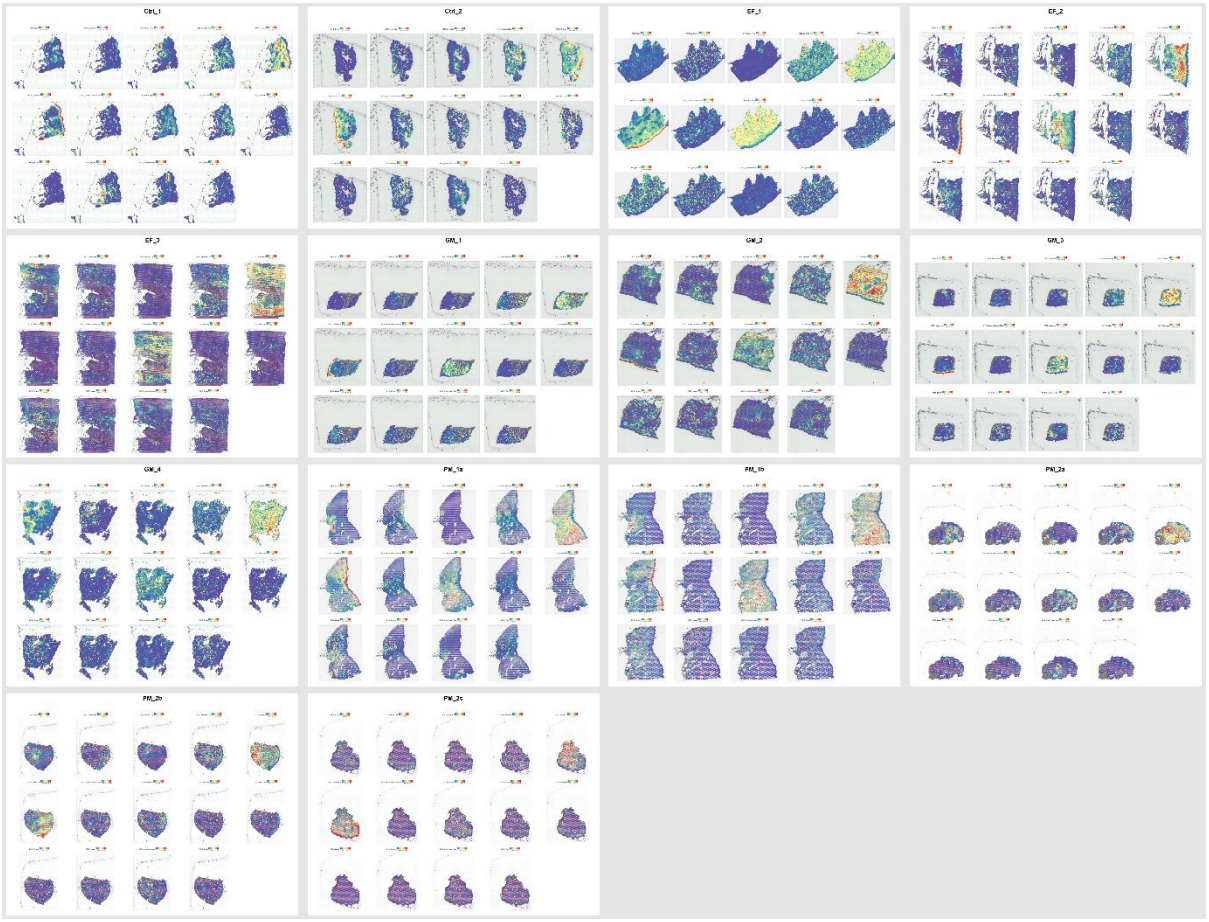
Supplementary Figure 3



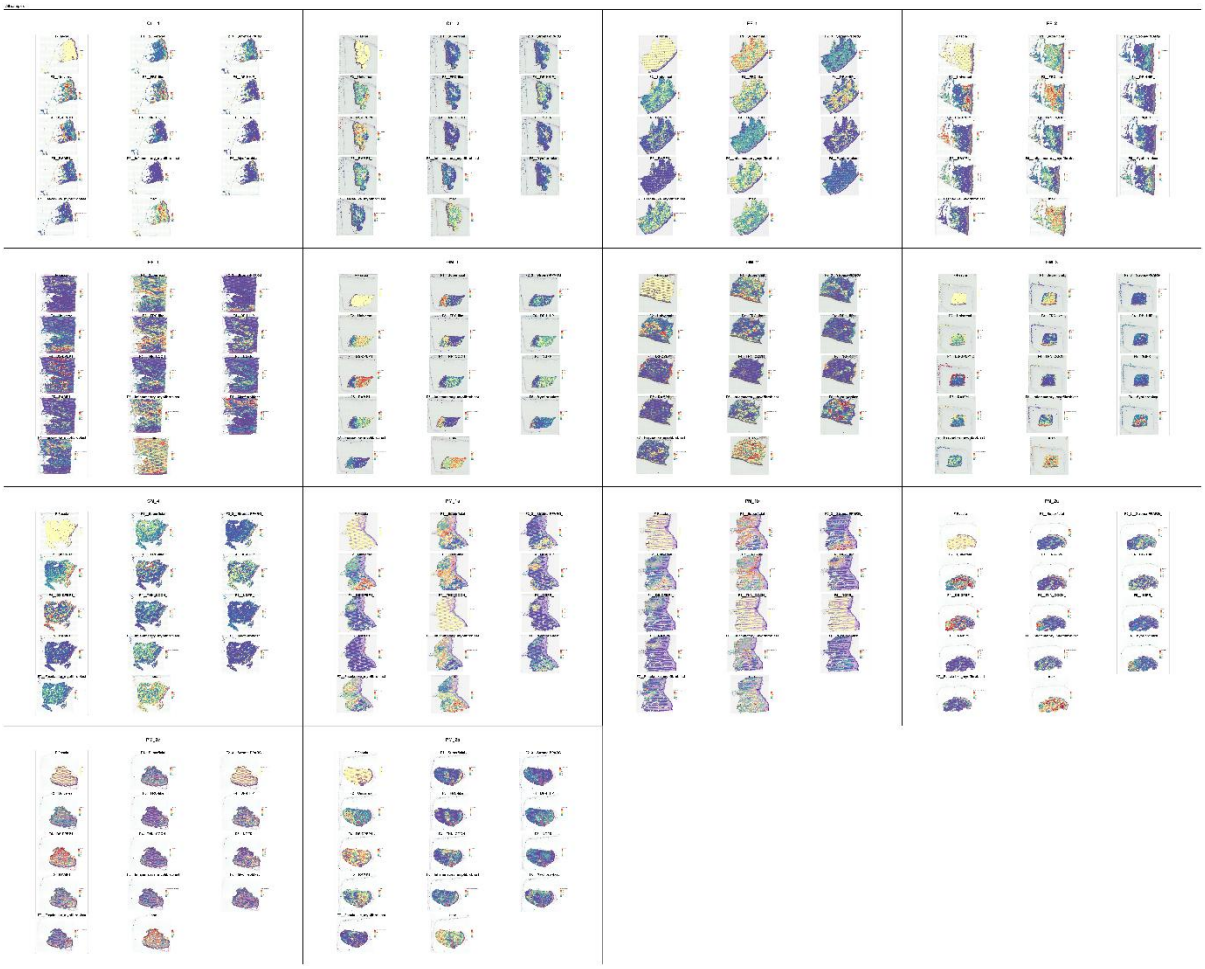
Supplementary Figure 4



Supplementary Figure 5



Supplementary Figure 6



Supplementary Figure 7

