Supplementary to:

Spatial transcriptomics reveals strong association between SFRP4 and extracellular matrix remodeling in prostate cancer

Maria K. Andersen[1], Sebastian Krossa[1], Elise Midtbust[1,2], Christine A. Pedersen[1], Maximillian Wess[1], Therese S. Høiem[1], Trond Viset[3], Øystein Størkersen[3], Ingunn Nervik[4], Elise Sandsmark[1,5], Helena Bertilsson[4,6], Guro F. Giskeødegård[7], Morten B. Rye[2,4,8,9] and May-Britt Tessem[1,2]

1. Department of Circulation and Medical Imaging, NTNU - Norwegian University of Science and Technology
2. Clinic of Surgery, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway
3. Department of Pathology, St. Olavs Hospital, Trondheim University Hospital
4. Department of Clinical and Molecular Medicine, NTNU - Norwegian University of Science and Technology
5. Department of Radiology and Nuclear Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway
6. Central Norway Regional Health Authority, Stjørdal, Norway.
7. Department of Public Health and Nursing, NTNU - Norwegian University of Science and Technology
8. Clinic of Laboratory Medicine, St.Olavs Hospital, Trondheim University Hospital, Trondheim, Norway
9. BioCore - Bioinformatics Core Facility, NTNU – Norwegian University of Science and Technology, Trondheim, Norway

Supplementary Methods

Proteomics on laser micro dissected (LMD) regions

Fresh frozen sections to be analyzed with MS proteomics were placed on sterile membrane slides (MMI, Eching, Germany). The sections were submerged in cold 70 % ethanol, dried for 30 s at room temperature, dehydrated in cold 100 % ethanol, air dried and stored at -80 °C.

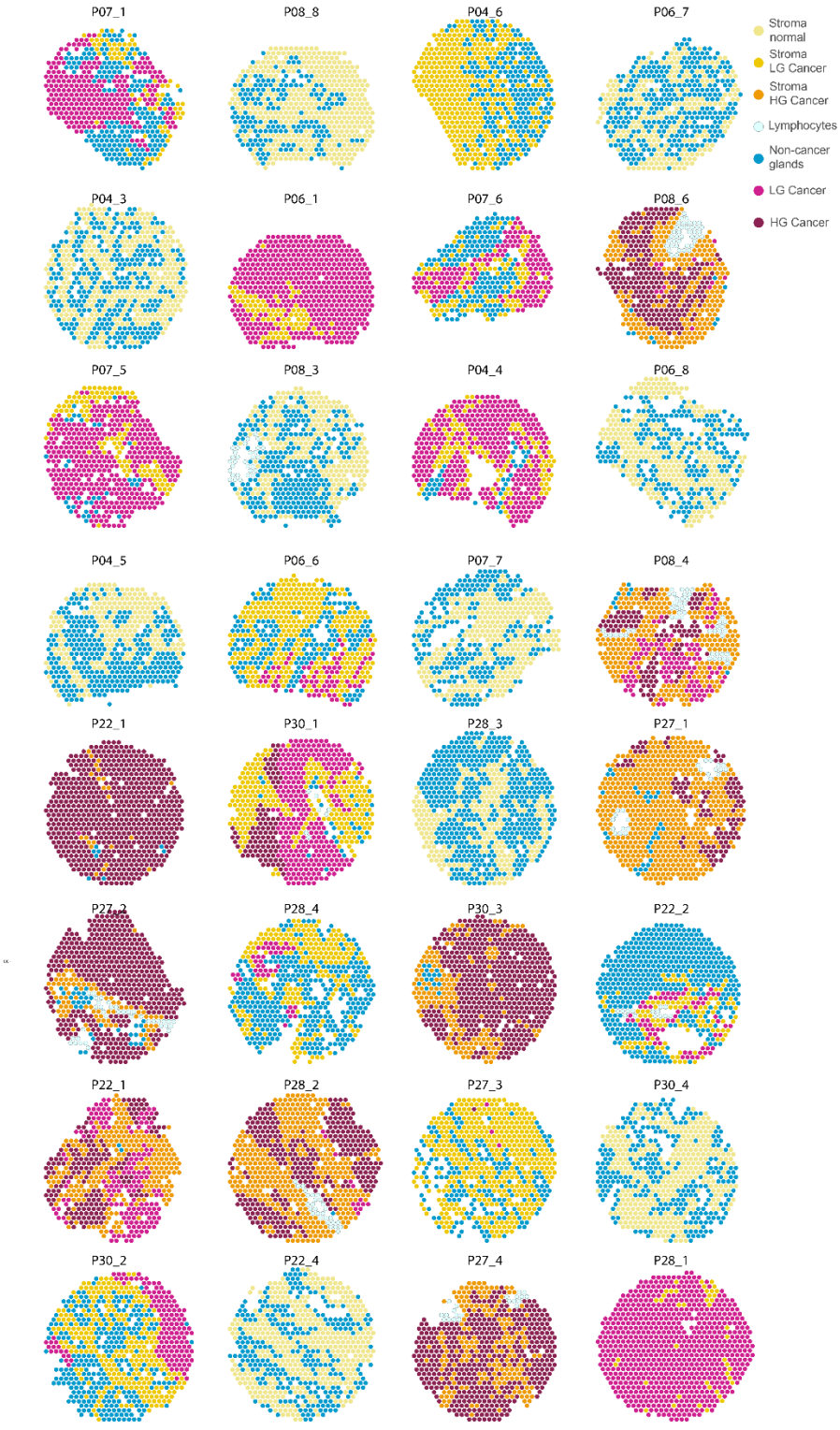
The tissue sections were rinsed with ultrapure water (15 s) and stained with hematoxylin (30 s), followed by ultrapure water (15 s) and dehydration by sequentially washing with 50 %, 70 %, 95 % (15 s each), and 100 % ethanol (60 s). Directly after staining, the membrane glass slides were digitally scanned, vacuum packed and stored in -80 °C until LMD.

LMD were performed using a Leica LMD 6000 system (Wetzlar, Germany). Prior to cutting, the membrane slides were defrosted and air dried. LMD areas were cut and collected in 20 µl TE buffer within 2.5 hours after drying, centrifuged, and subsequently placed in -80 °C storage. A total of 114 areas were cut, including (but not limited to) 28 stroma regions, 30 normal gland regions and 37 cancer glands regions. Area of regions cut was in the range 39349 - 4165184 μm^2, while number of cells varied from 187 to 9171.

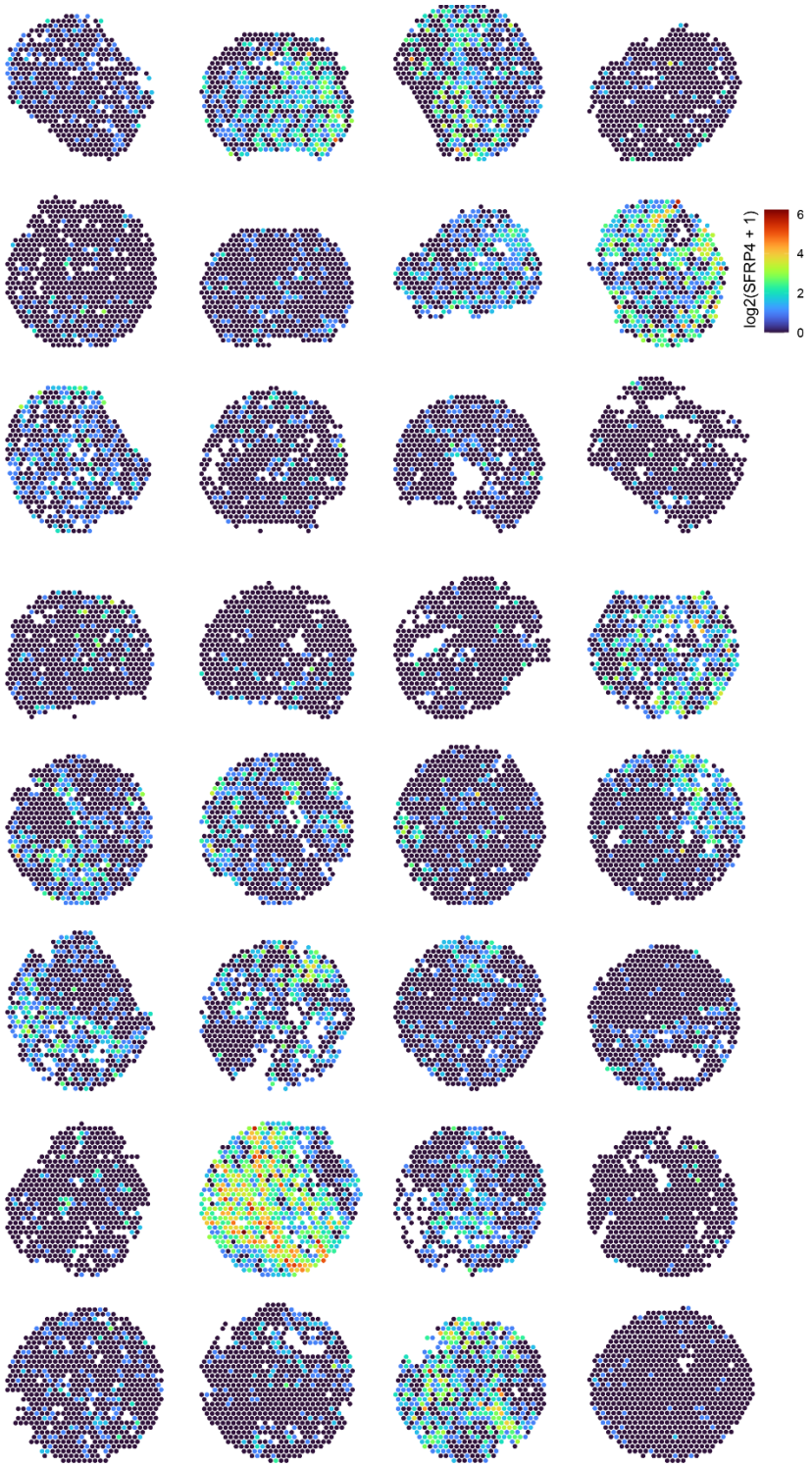
The LMD tissue areas were thawed, resuspended in 20 µl of lysis buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA, 0.02 % Zwittergent 3-16), heated for 60 min at 98 °C, and sonicated for 30 min. Iodoacetamide (IAM; 15 mM, 15 min in the dark at room temperature) was used to alkylate cysteines after reduction of disulfide bridges (by dithiothreitol (DTT, 10 mM, 56 °C for 30 min)). The proteins were cleaved into peptides by overnight tryptic digestion (0.03 µg, 37 °C). Tryptic activity was terminated by acetic acid (HAC; final concentration of 0.5 %) and desalted with in-house packed C18 Stagetips (Empore, 3M). For LC-MS we used a TimsTOF Pro2 instrument coupled in-line with a nanoElute ultraperformance liquid chromatography (UPLC) system (Bruker Daltonics). Samples were analyzed in a randomized order, and

approximately every tenth run an aliquoted prostate tissue standard (created from 5 different sections) were measured. Liquid chromatography was performed using a Pepsep C18 column (25 cm×150 µm, 1.5 µm; Bruker Daltonics) with a gradient of 0-37 % acetonitrile, 0.1 % formic acid for 75 min. The timsTof instrument was operated in the DDA PASEF mode with 10 PASEF scans per acquisition cycle and accumulation and ramp times of 100 ms each. The 'target value' was set to 20,000 and dynamic exclusion was activated and set to 0.4 min. The quadrupole isolation width was set to 2 Th for $m/z < 700$ and 3 Th for $m/z > 800$.

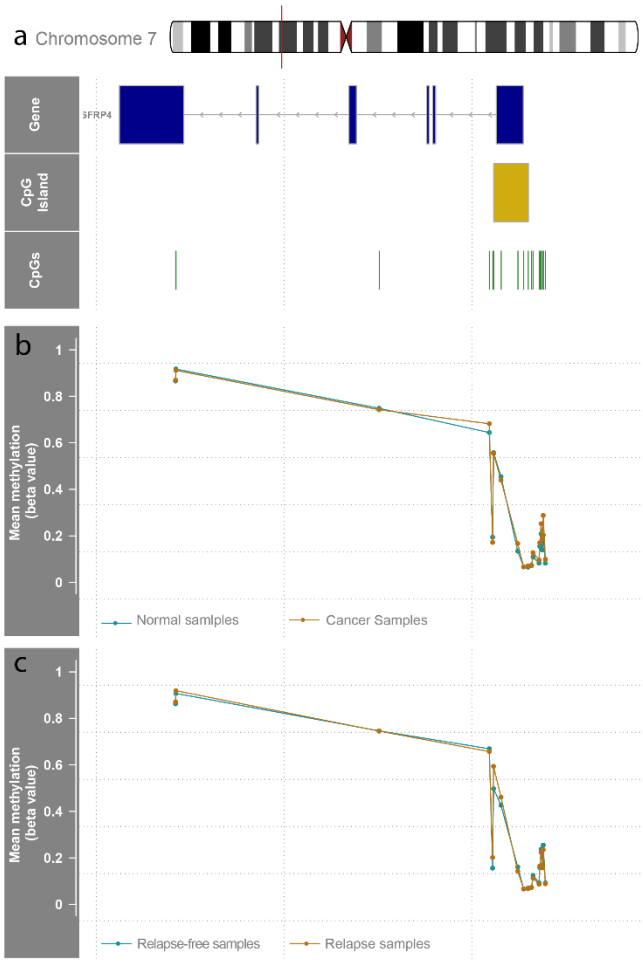
Proteins were identified using MaxQuant (version 2.0.3.1). Trypsin was chosen as proteolytic enzyme (max two missing cleavages). Carbamidomethylation of cysteine residues was set as fixed modifications and oxidation of methionine, acetylation of protein N-terminus, and deamidation of asparagine/glutamine was chosen as dynamic post-translational modifications. Each run in MaxQuant utilized *m/z* and retention time (RT)-values and aligned them to each of the samples (window: one minute match-between-run function and 20 min overall sliding window) using a clustering-based technique. The queries (including protein isoforms) were searched against the Human proteome database from Uniprot (downloaded June 2022; https://www.uniprot.org/proteomes/UP000005640). Contaminants were searched against MaxQuant's internal contaminants database using inbuilt Andromeda. Maximum false discovery rate was set to 1 % (FDR <0.01) for both peptides and proteins, where unique peptides with high confidence were used for final protein identification.



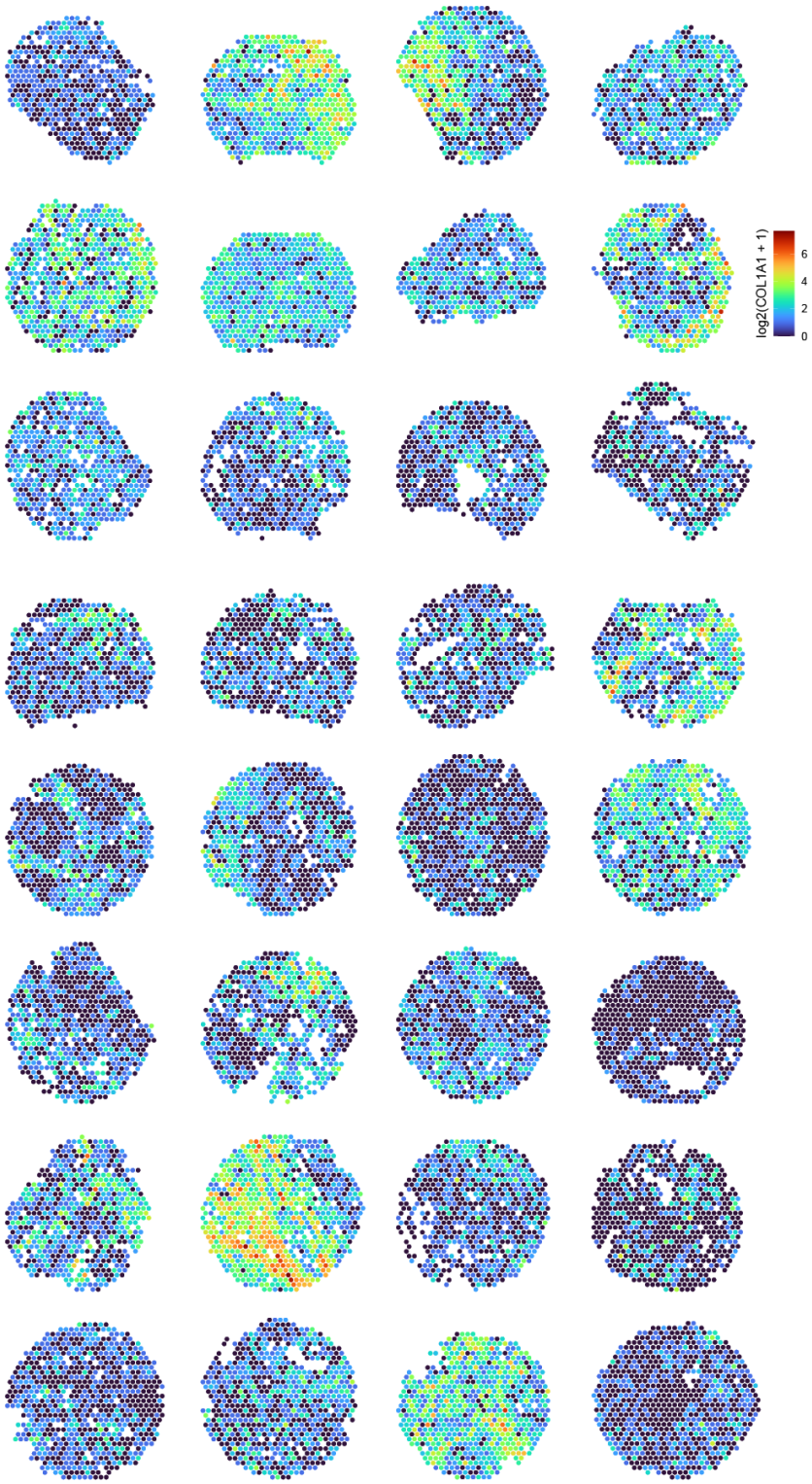
Supplementary Figure S1: Histopathology classifications for each spatial transcriptomics spot. Low grade (LG) cancer included Gleason grade group 1 and 2, while high grade (HG) cancer represents grade groups 3 through 5. Stroma spots were assigned according to if they coappeared in samples with normal glands, LG or HG cancer glands.



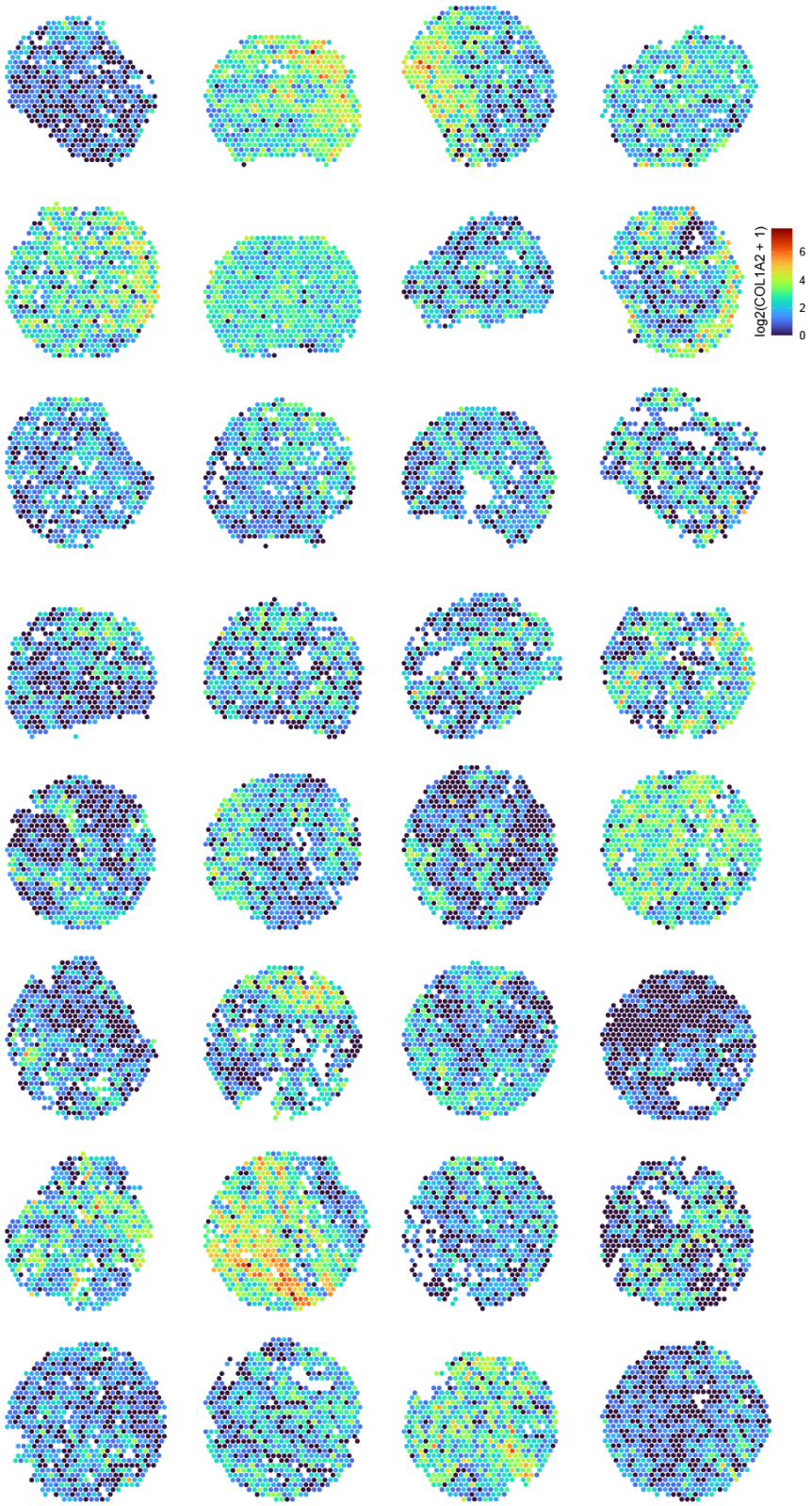
Supplementary Figure S2: Spatial gene expression distribution of *SFRP4*. Gene counts are cell count normalized and $\log_2$-transformed.



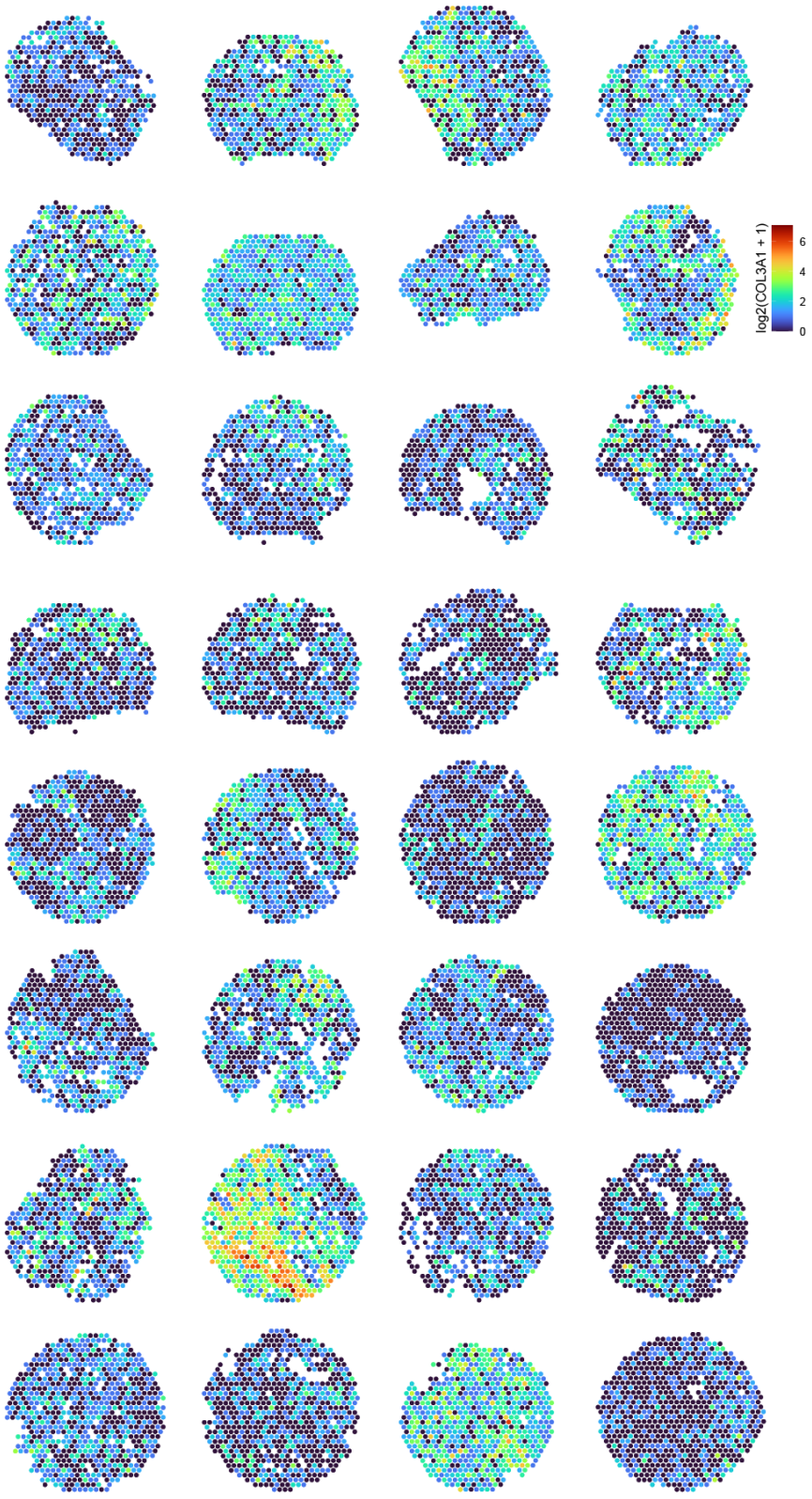
Supplementary Figure S3: *SFRP4* DNA Methylation. a) Diagram shows chromosome and genome location of the *SFRP4* gene. Mean methylation values are presented for b) comparing normal to cancer samples and c) comparing control to relapse samples.



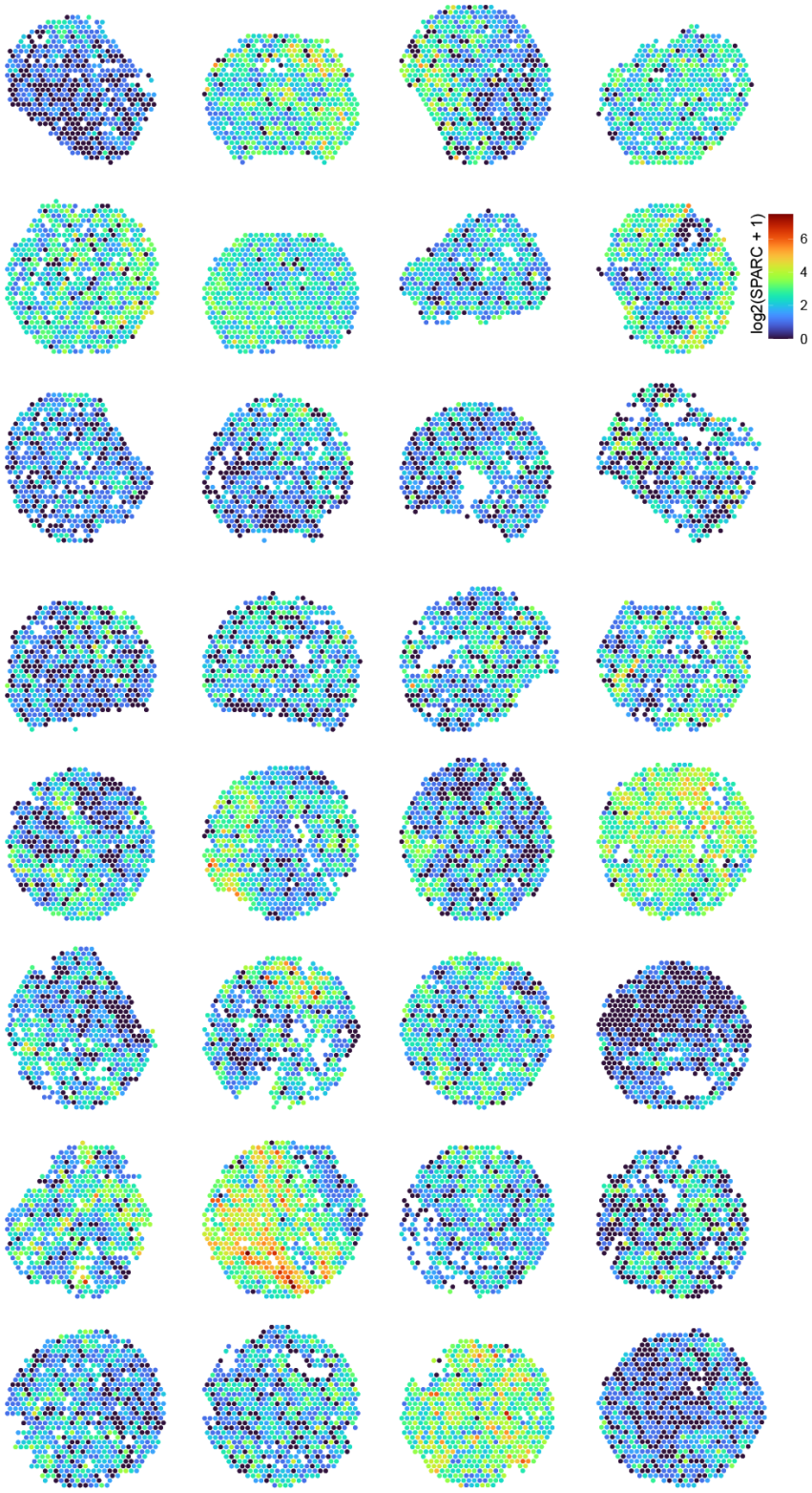
Supplementary Figure S4: Spatial gene expression distribution of *COL1A1*. Gene counts are cell count normalized and $\log_2$-transformed.



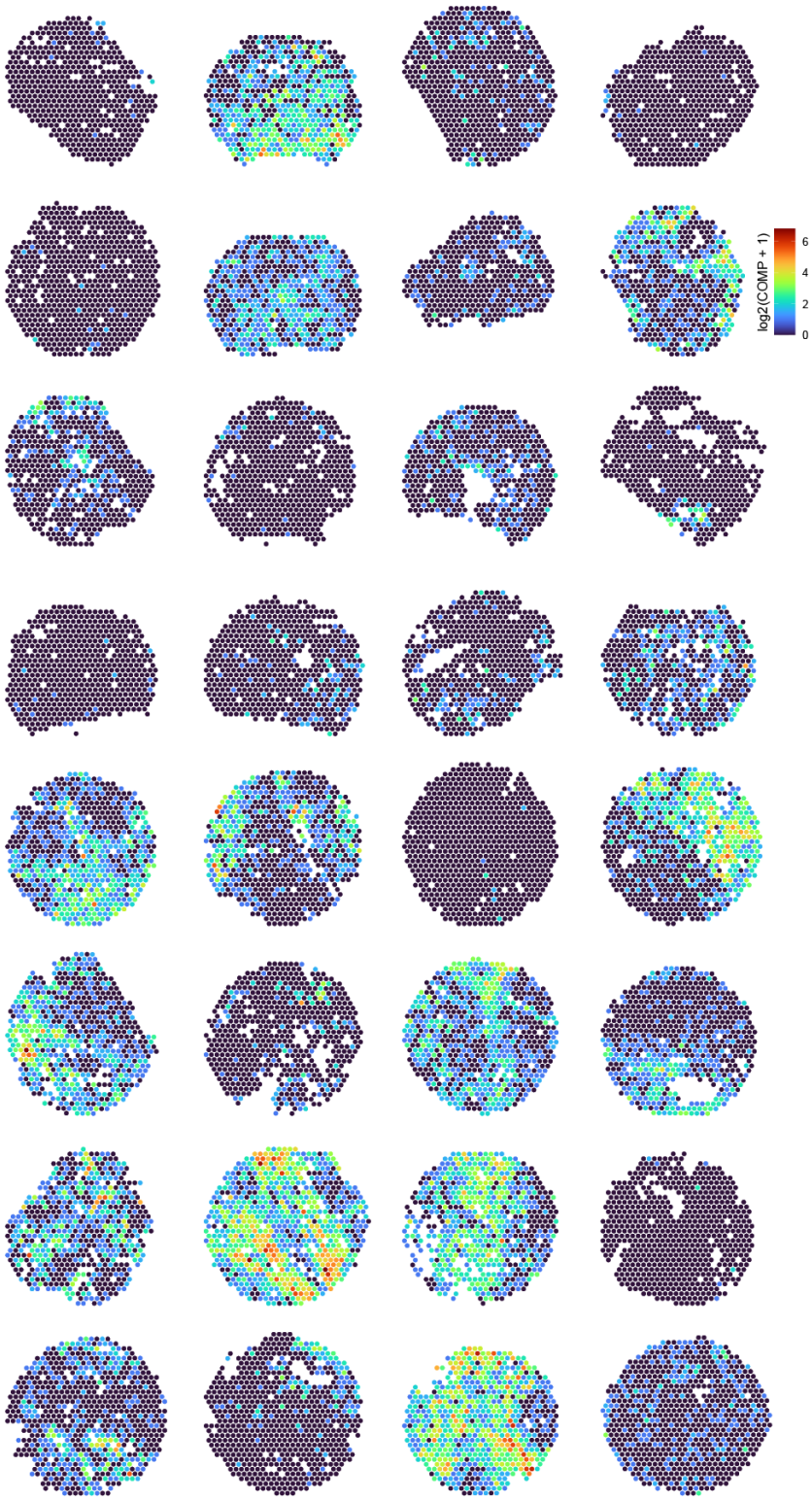
Supplementary Figure S5: Spatial gene expression distribution of *COL1A2*. Gene counts are cell count normalized and $\log_2$-transformed.



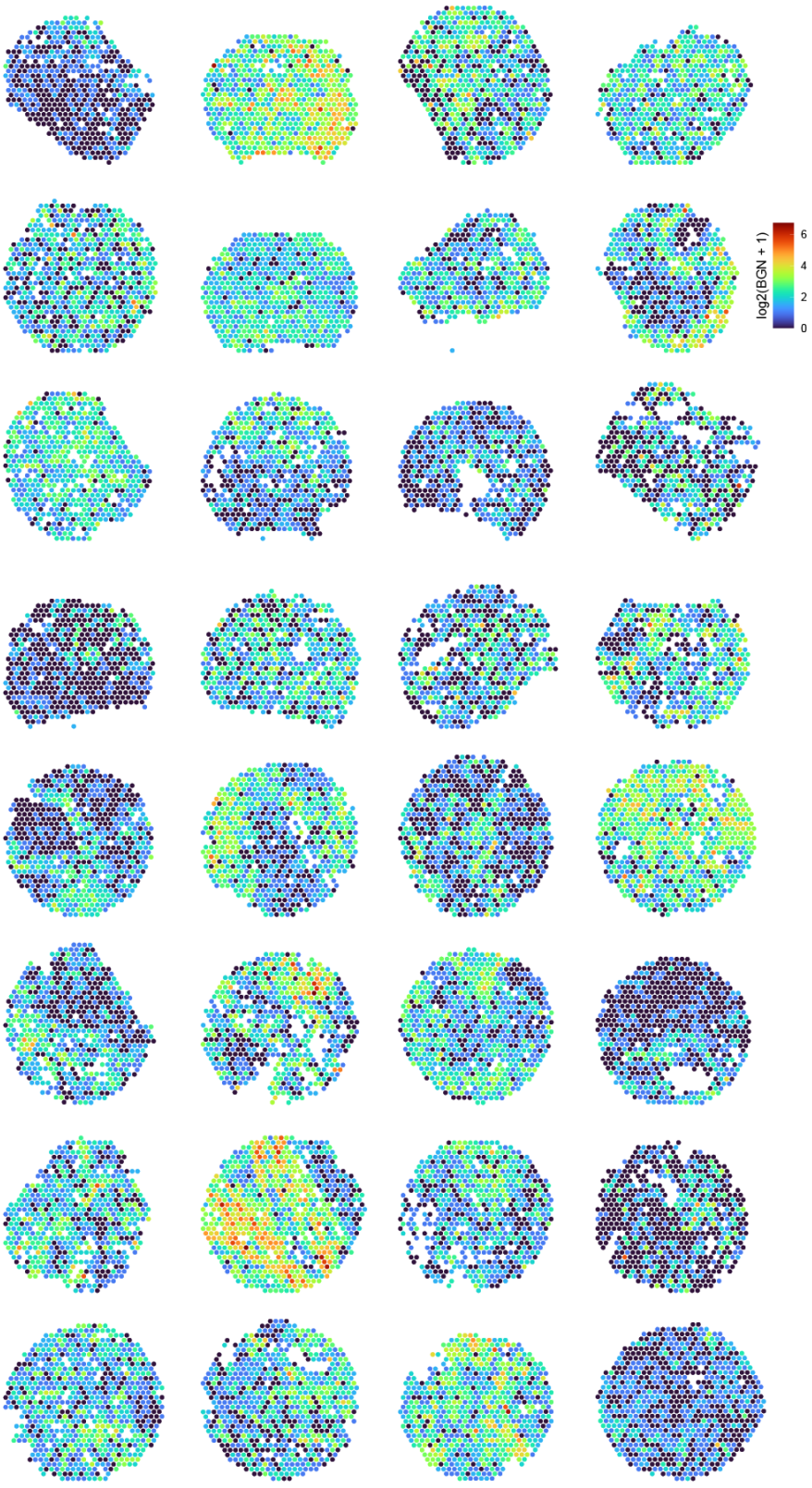
Supplementary Figure S6: Spatial gene expression distribution of *COL3A1*. Gene counts are cell count normalized and $\log_2$-transformed.



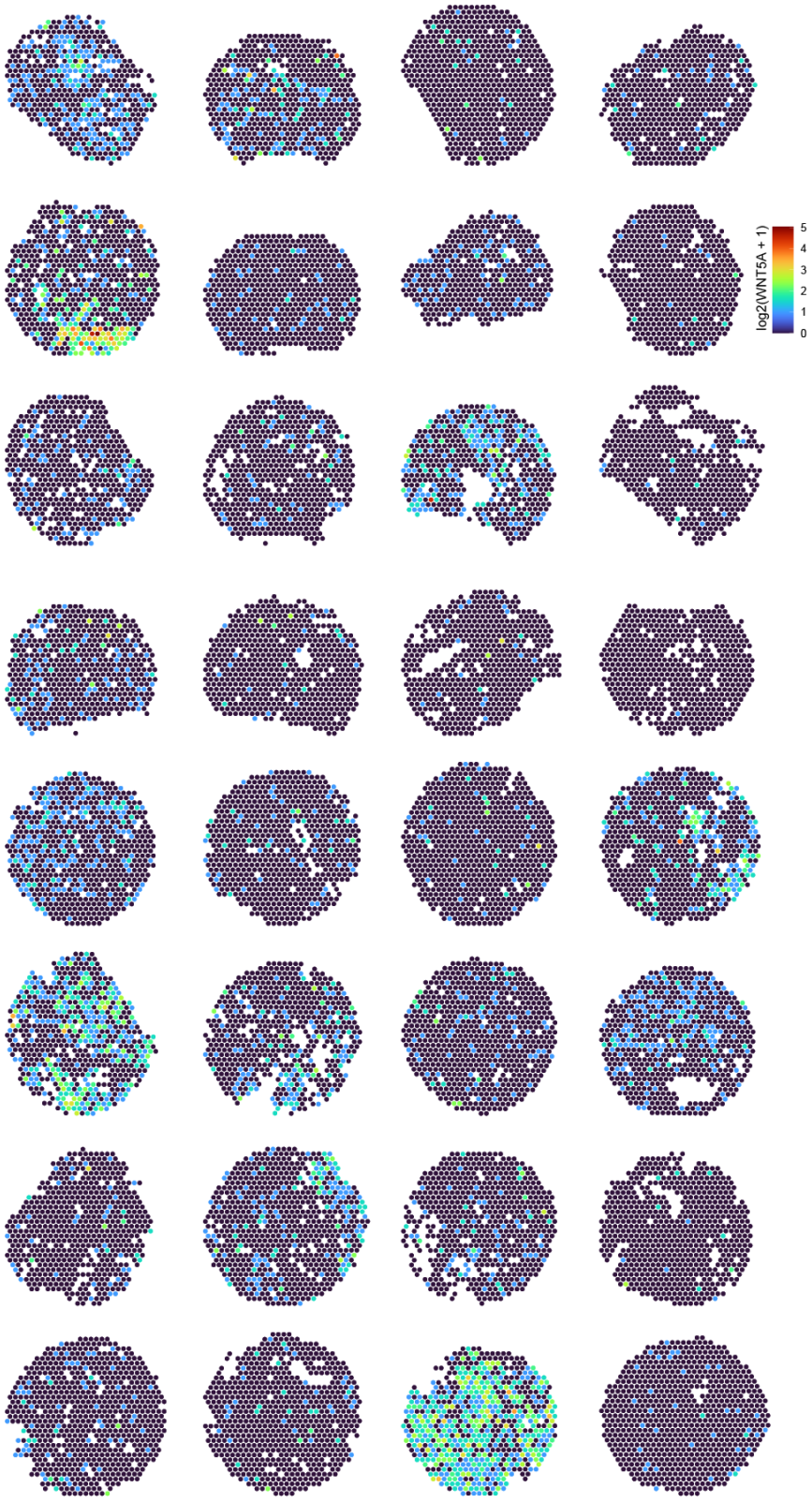
Supplementary Figure S7: Spatial gene expression distribution of *SPARC*. Gene counts are cell count normalized and $\log_2$-transformed.



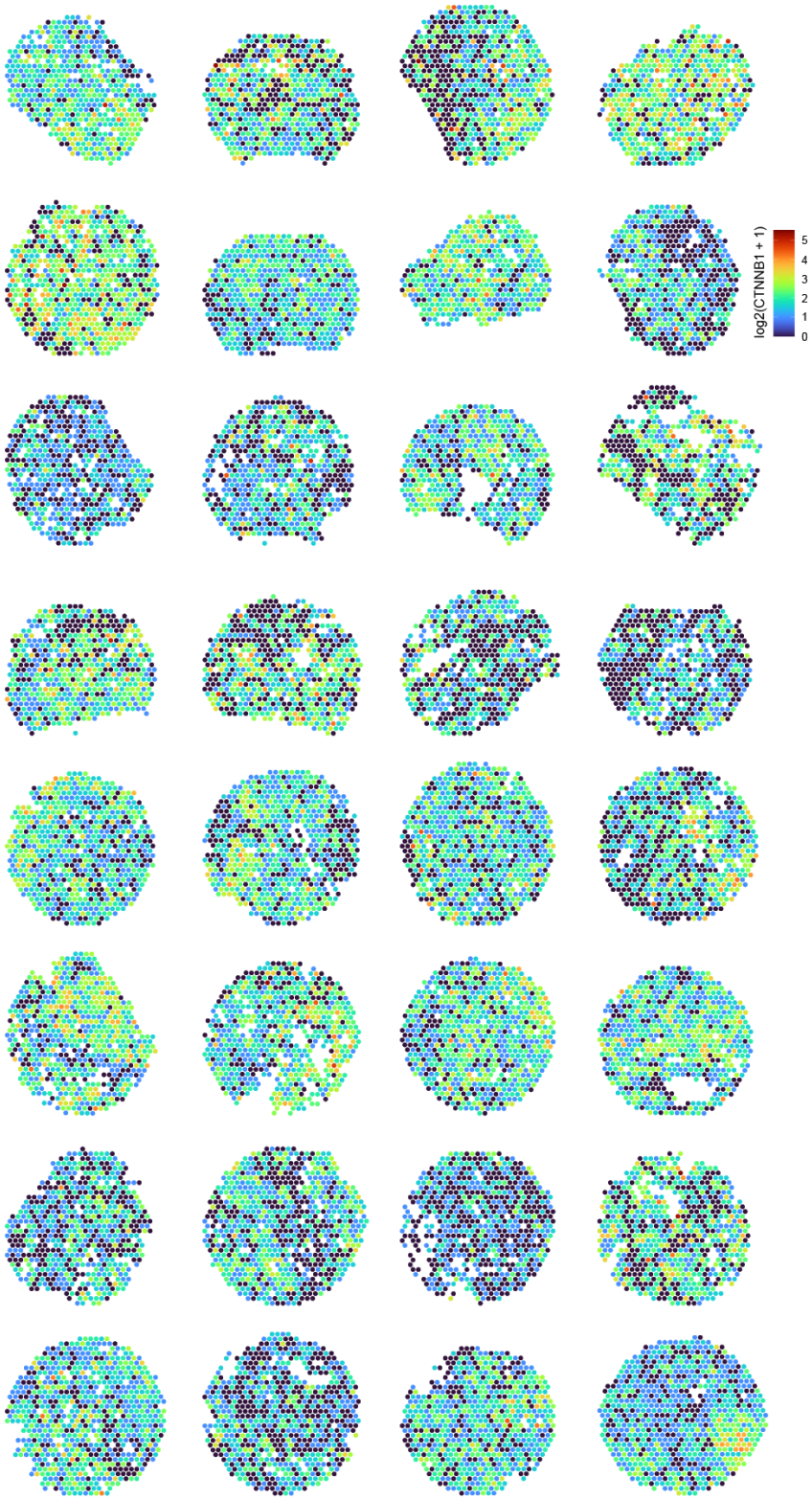
Supplementary Figure S8: Spatial gene expression distribution of *COMP*. Gene counts are cell count normalized and $\log_2$-transformed.



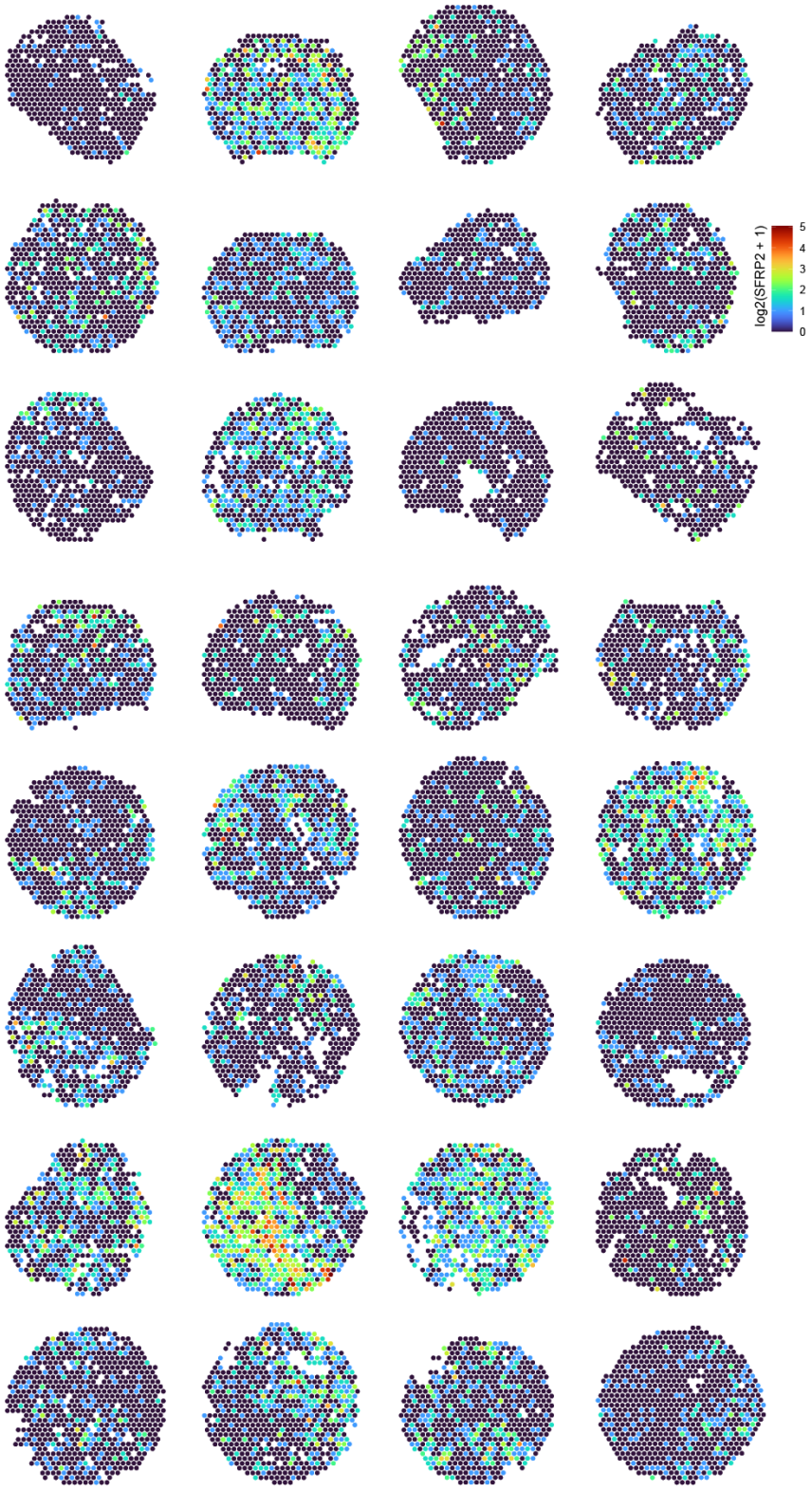
Supplementary Figure S9: Spatial gene expression distribution of *BGN*. Gene counts are cell count normalized and $\log_2$-transformed.



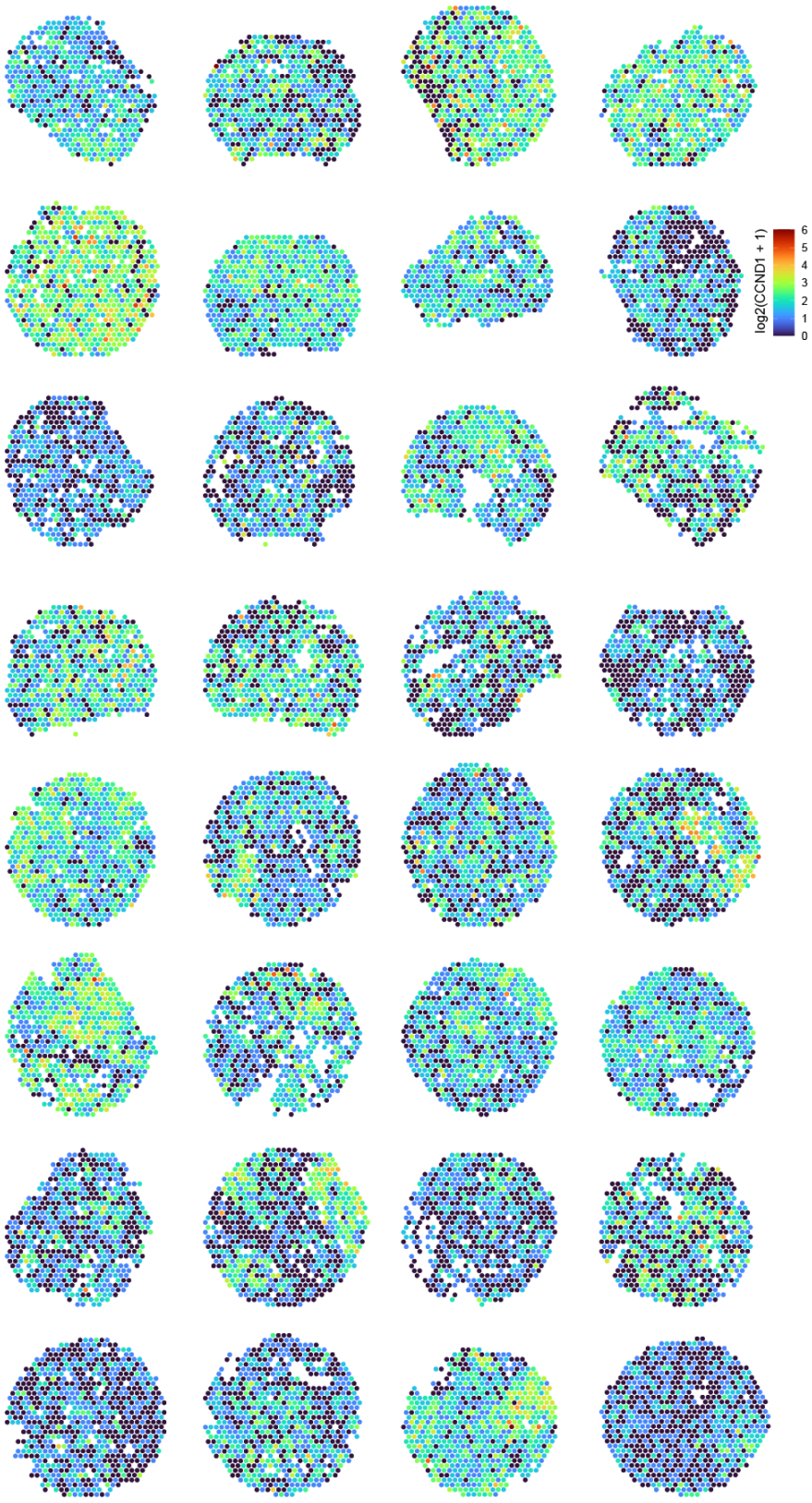
Supplementary Figure S10: Spatial gene expression distribution of *WNT5A*. Gene counts are cell count normalized and $\log_2$-transformed.



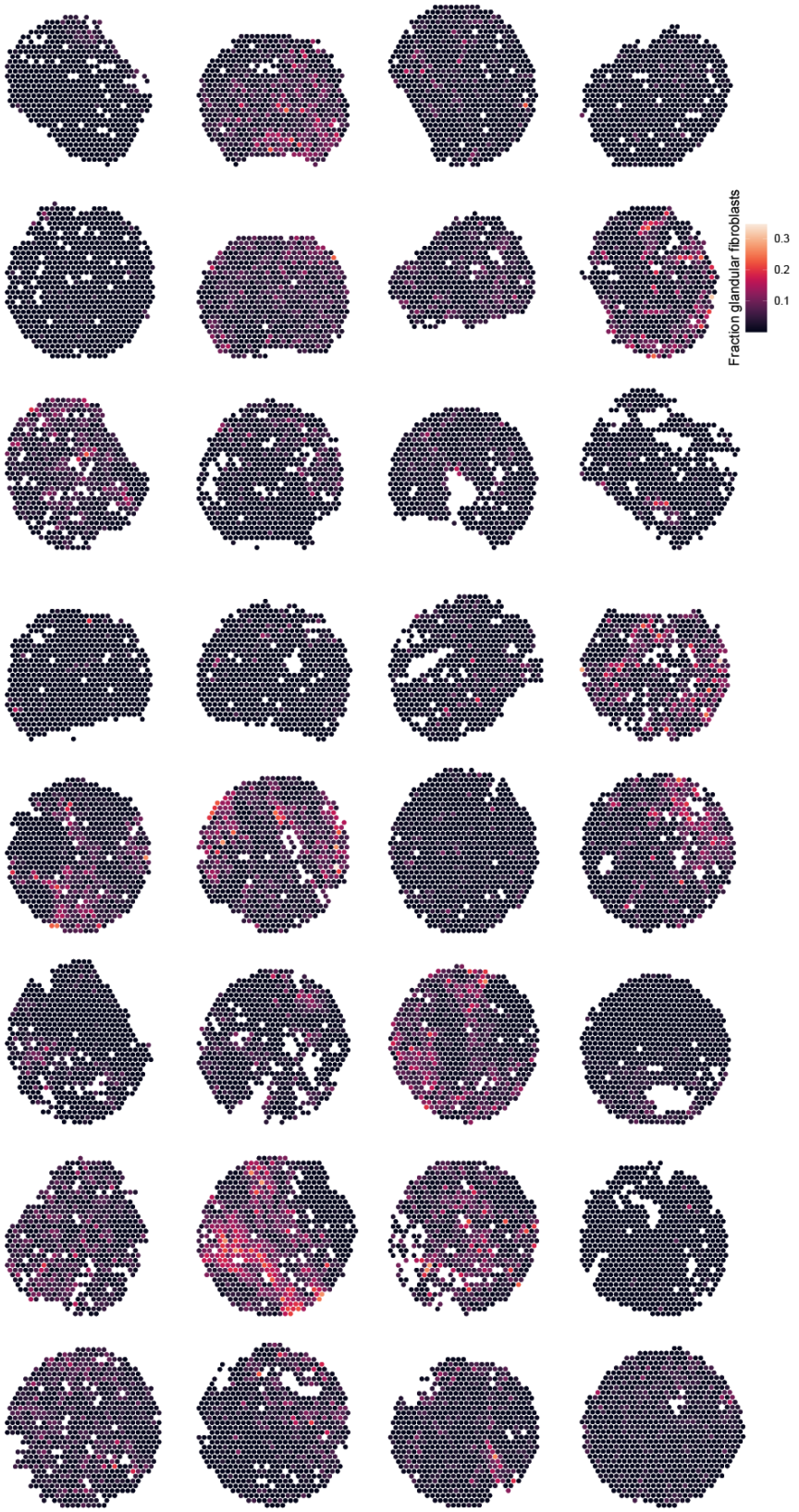
Supplementary Figure S11: Spatial gene expression distribution of *CNNB1*. Gene counts are cell count normalized and $\log_2$-transformed.



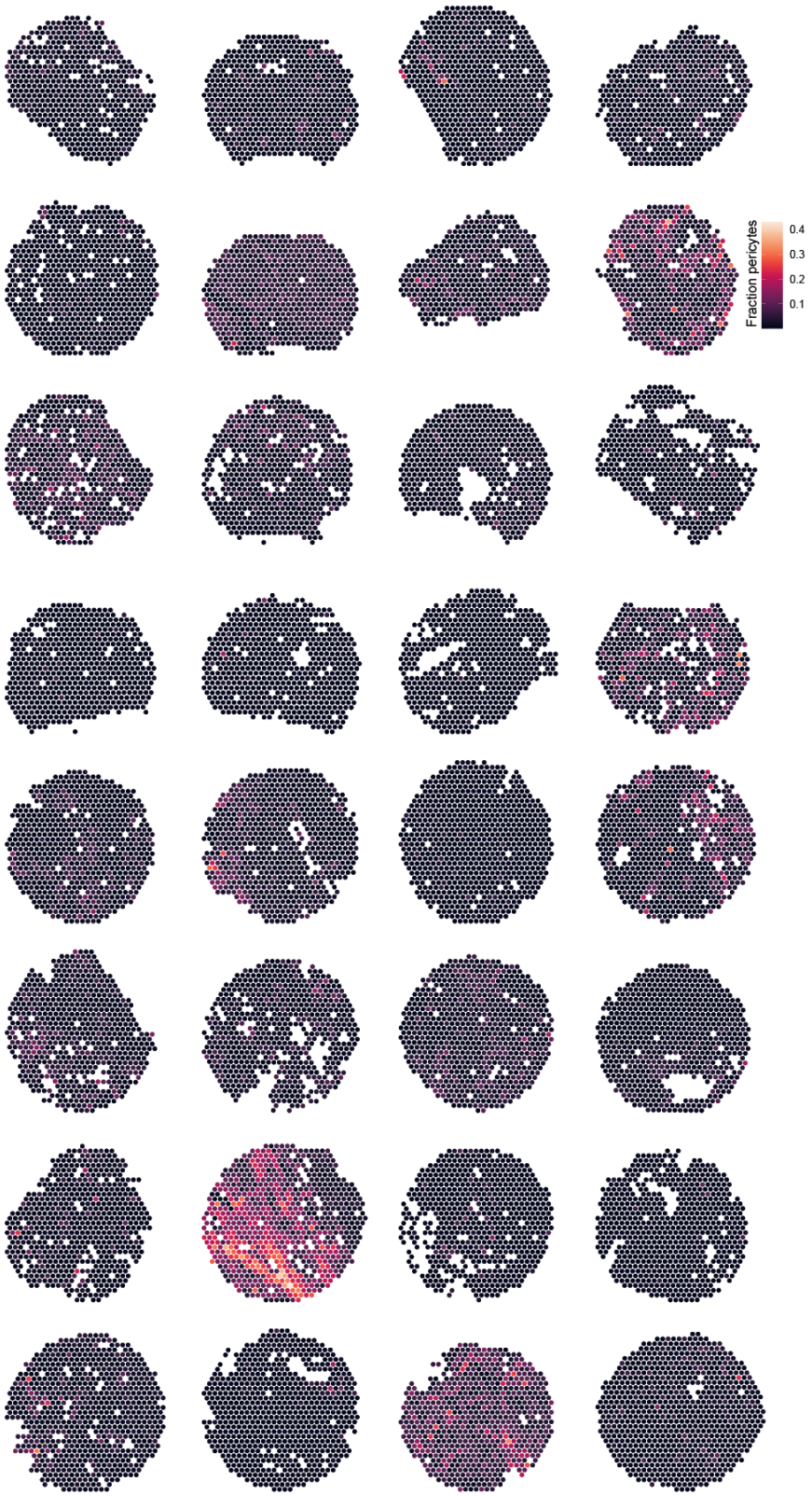
Supplementary Figure S12: Spatial gene expression distribution of *SFRP2*. Gene counts are cell count normalized and $\log_2$-transformed.



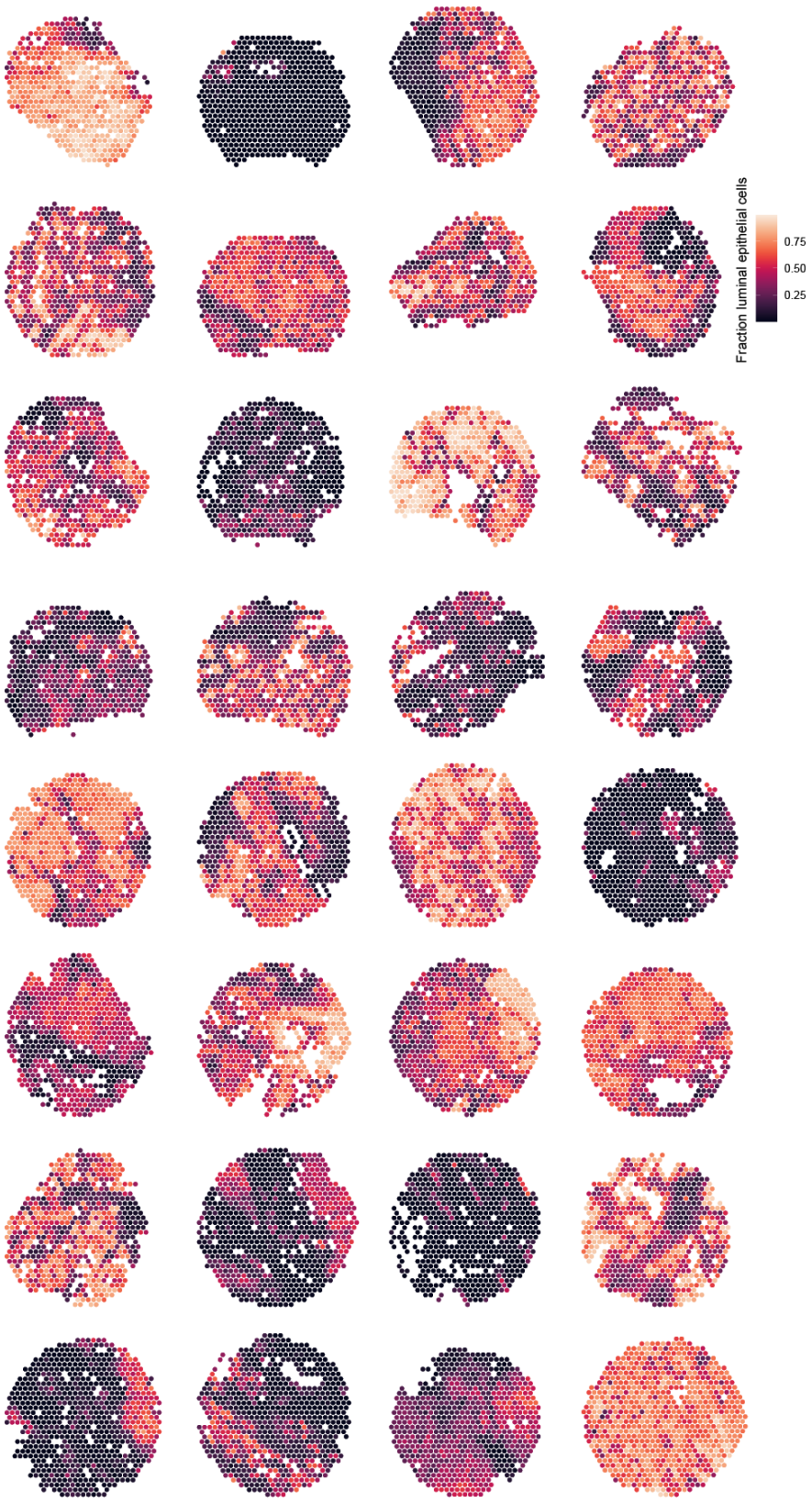
Supplementary Figure S13: Spatial gene expression distribution of *CCND1*. Gene counts are cell count normalized and $\log_2$-transformed.



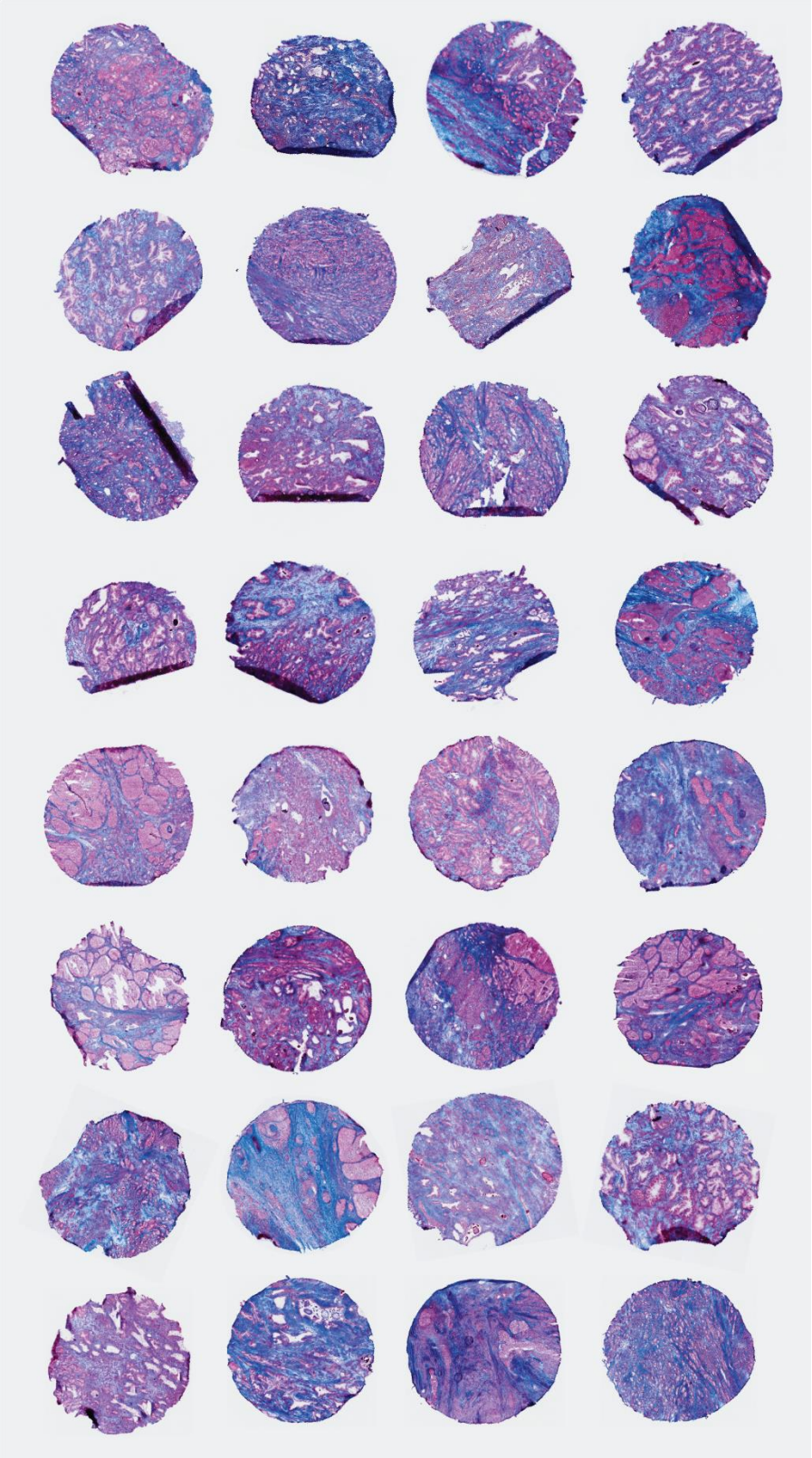
Supplementary Figure S14: Estimated cell fraction of glandular fibroblasts in each spatial transcriptomics spot.



Supplementary Figure S15: Estimated cell fraction of pericyte in each spatial transcriptomics spot.



Supplementary Figure S16: Estimated cell fraction of luminal epithelial cells in each spatial transcriptomics spot.



Supplementary Figure S17: Masson's trichrome stain. Staining was performed on serial sections of the same samples used for spatial transcriptomics.